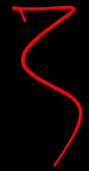
CS4.301 Data & Applications

Ponnurangam Kumaraguru (“PK”)
#ProfGiri @ IIIT Hyderabad



pk.profgiri /in/ponguru @ponguru Ponnurangam.kumaraguru

Protocol

Who am I?

- ~~Assistant~~ ~~Associate~~ Professor of Computer Science
- Ph.D. from School of Computer Science, Carnegie Mellon University (CMU)
- Research interests
  - Computational Social Science
  - Social (Societal) Computing
  - Privacy & Security in Social Media
- Courses I teach
  - Data & Applications (2), 4+
  - Responsible & Safe AI (1), 4+
  - Online Privacy (1)
  - Privacy and Security in Online Social Media (8), 4+
  - Designing Human Centered Systems (5), 4+
  - Research methods / Advanced research methods (2), 4+
  - Foundations of Computer Security (5), 4+
  - Big Data & Policing (1), 4+

Who you are?

CND

ECD

EHD

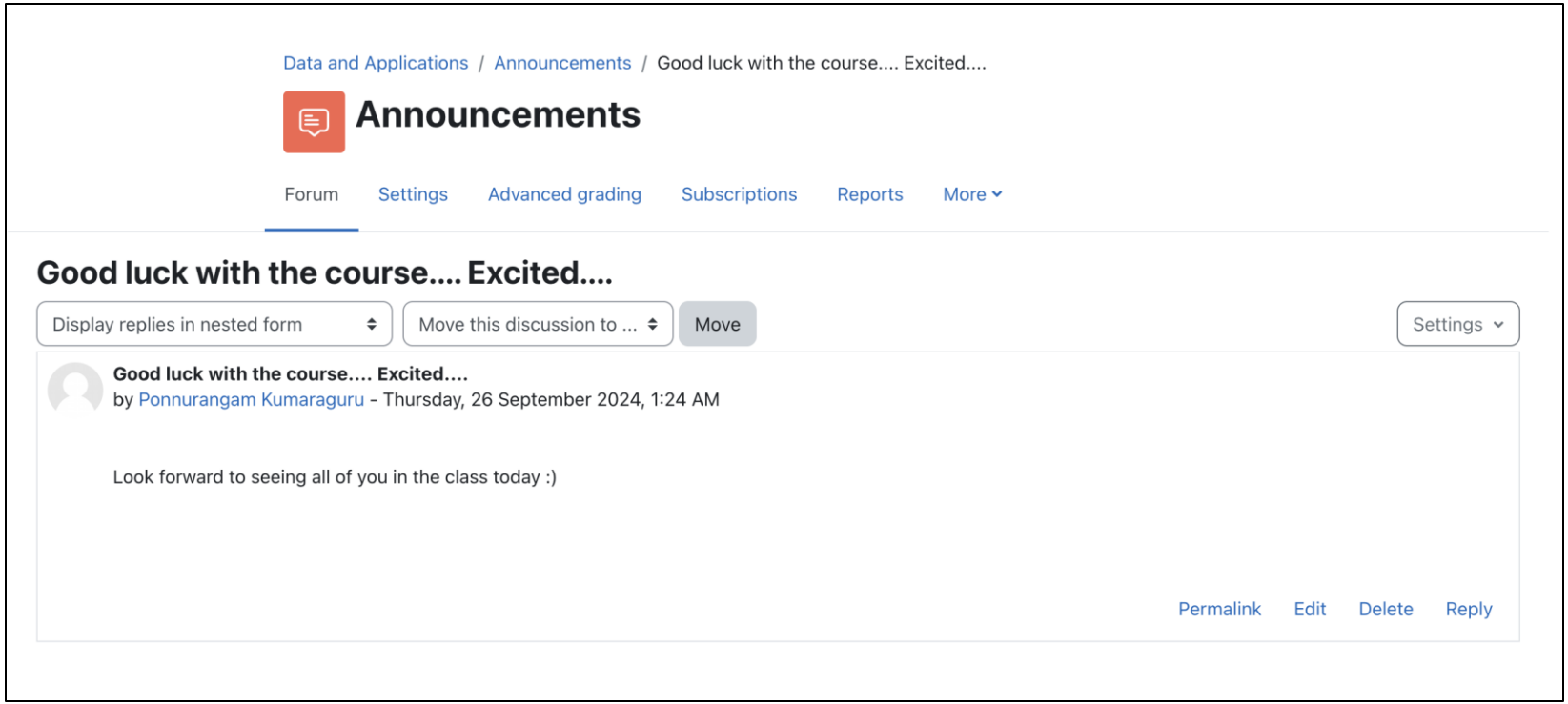
CSD

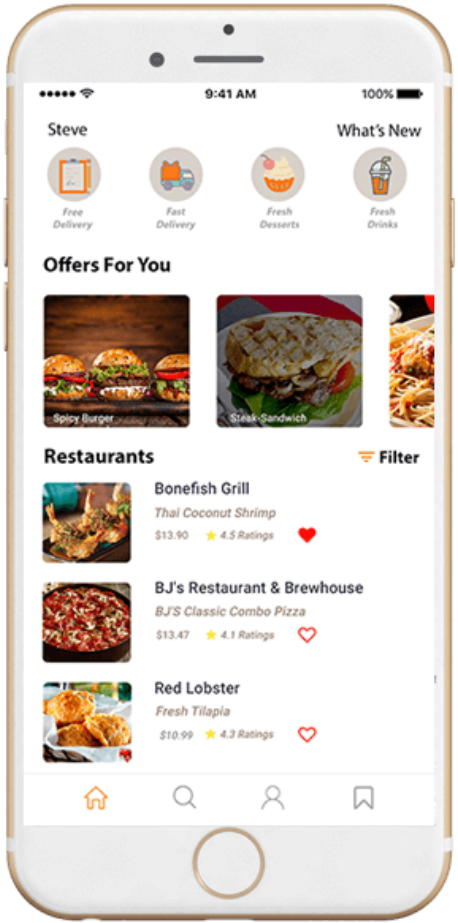
ECE

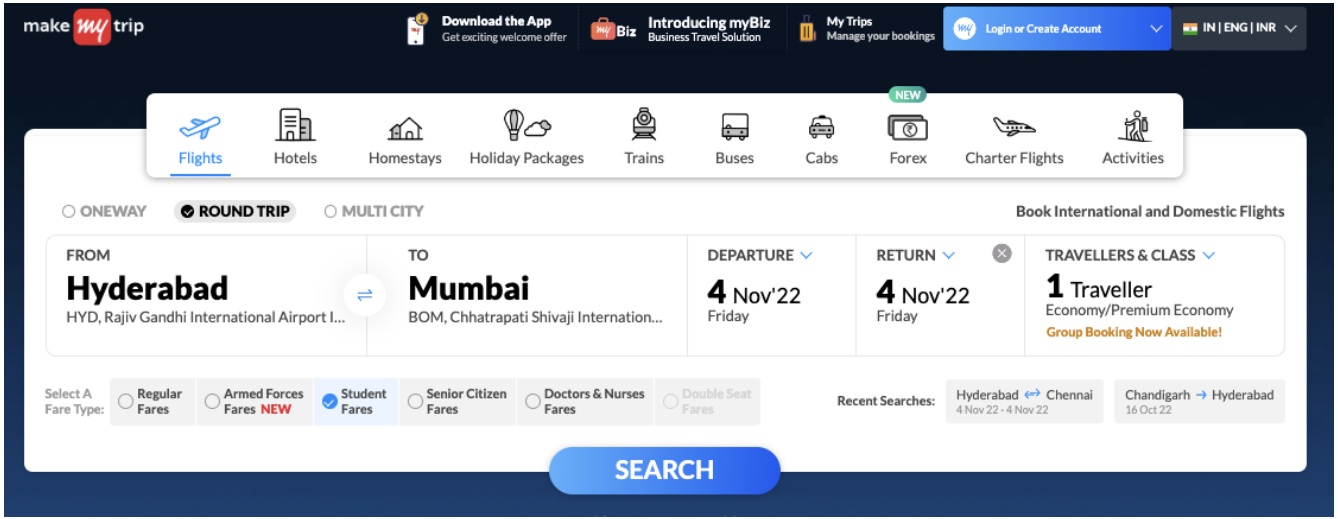
CLD

CSE

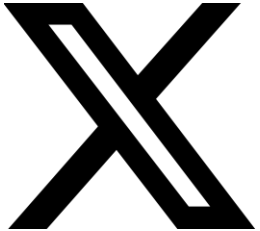
???


Data and Applications / Announcements / Good luck with the course.... Excited....
Announcements
Forum
Settings
Advanced grading
Subscriptions
Reports
More
Good luck with the course.... Excited....
Display replies in nested form
Move this discussion to ...
Move
Settings
Good luck with the course.... Excited....
by Ponnurangam Kumaraguru - Thursday, 26 September 2024, 1:24 AM
Look forward to seeing all of you in the class today :)
Permalink
Edit
Delete
Reply





What kind of data are we generating here?



What kind of data are we generating here?

Grading, Relative

Type of Evaluation	Weightage (in %)
Class Quizzes (3)	15
Assignments / Homeworks (4)	20
Mid sem exam (Quiz-2 as scheduled in almanac)	15
Project	30
End Sem Exam	20

TAs

12 TAs

Students will be assisged among TAs for all evaluations

Aditya Mishra
Anika Roy
Anirudh Vempati
Anish R Joishy
Chetan Mahipal
Debangan Mishra
Devika Umesh Bej
Jain Hemang Ashok
Pratishtha Saxena
Priyanshul Govil
Rohan Kumar
Shailender Goyal
Tejas Cavale

Memory challenge for me ☺

Plagiarism

What is it?

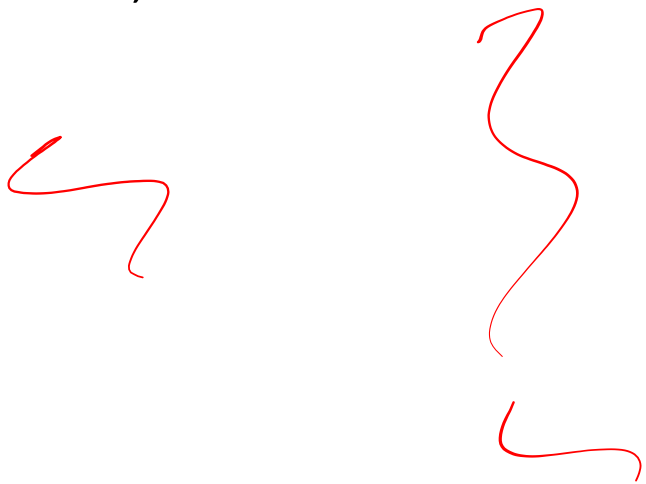
Copying HWs

Any content taken from another source without citation

Whatever policy from IIITH

Moodle

We will use Moodle for all content sharing – slides, HWs, announcements, clarifications, etc.



Service Level Agreement

Any question / clarification ask, if not urgent, will be answered in 24 hrs

If anything urgent, feel free to attach the time in which you want the answer, we will try to respond

TAs are your 1st point of contact only on escalation, you will bring it up to me

Please do not email only me ☹

✗ Do not Send WA msg to TAs.

Topics that we will cover

Relational Database Systems

SQL

Database design process

Data Models, Normalization

Soups

Cream of Tomato	165
Veg Clear Soup	165
Veg Hot & Sour Soup	165
Veg Corn Soup	165
Veg Silver Soup	165
Veg Cantonese	165
Veg Manchow	165

Starters - Chinese

Crispy Vegetable	300
Veg. Gold Coin	300
Veg. Manchurian	335
Veg. Spring Roll	335
Gobi Manchurian	335
Chutneys Spl. Spring Roll	335
Chilly Mushroom	335
Mushroom Manchurian	335
Diced Paneer Red Pepper	335
Baby Corn Manchurian	335
Hong Kong Mushroom	335
Crispy Babycorn	335
Crispy Corn	335
Chilly Paneer	335
Paneer/Gobi/Aloo 65	335
Paneer Majestic	335
Chilly Mushroom	335

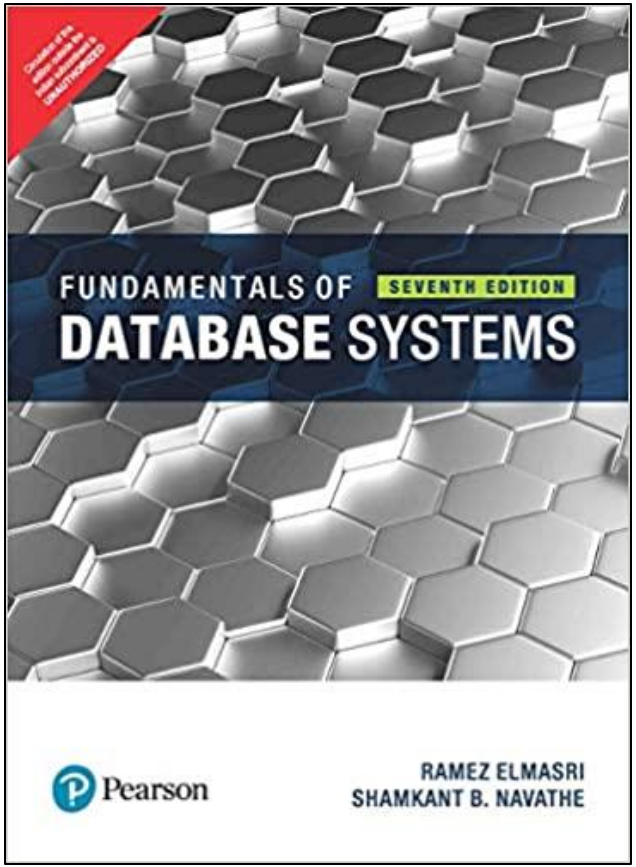
https://www.dineout.co.in/hyderabad/chutneys-madhapur-west-hyderabad-11747/menu

Lecture # from	Class	Tutorial	Date	Day	Time	Topic	HW @ 23:59hrs	Quiz	Project
15	1		26-Sep	Th	1535 - 1700	Intro & data models			
16	2		30-Sep	M	1535 - 1700	Intro & data models	HW1 publish		
17	3		3-Oct	Th	1535 - 1700	DB design & ER models			
18	4		7-Oct	M	1535 - 1700	DB design & ER models		Quiz 1	
		1	9-Oct	W	0830 - 0955	Data models, DB Design, ER Models & project	HW2 publish		
19	5		10-Oct	Th	1535 - 1700	Relational DB	HW1 submission		
			12-Oct						Data requirements
20	6		14-Oct	M	1535 - 1700	Relational DB			
		2	16-Oct	W	0830 - 0955	Data models, DB Design, ER Models & project			
21	7		17 - 19 Oct			Institute Quiz			
22	8		21-Oct	M	1535 - 1700	Normalization	HW3 publish		
		3	23-Oct	W	0830 - 0955	ER models	HW2 submission		
23	9		24-Oct	Th	1535 - 1700	Normalization			ER Model
24	10		28-Oct	M	1535 - 1700	Normalization		Quiz 2	

		4	29-Oct	W	0830 - 0955	Relational DB			
		5	6-Nov	W	0830 - 0955	Normalization	HW3 submission		Relational Database design
25	10		7-Nov	Th	1535 - 1700	Normalization	HW4 publish		
26	11		11-Nov	M	1535 - 1700	SQL			
		6	13-Nov	W	0830 - 0955	SQL			Application
27	12		14-Nov	Th	1535 - 1700	SQL		Quiz 3	
28	13		18-Nov	M	1535 - 1700	SQL	HW4 submission		
		7	20-Nov	W	0830 - 0955	SQL + Revision			
			20-Nov			All marks check			
			22 - 29 Nov			End Sem			
			22 - 29 Nov			Final project demo			
			1 - 3 Dec			End Sem paper check			
			5-Dec			Grades to be submitted			

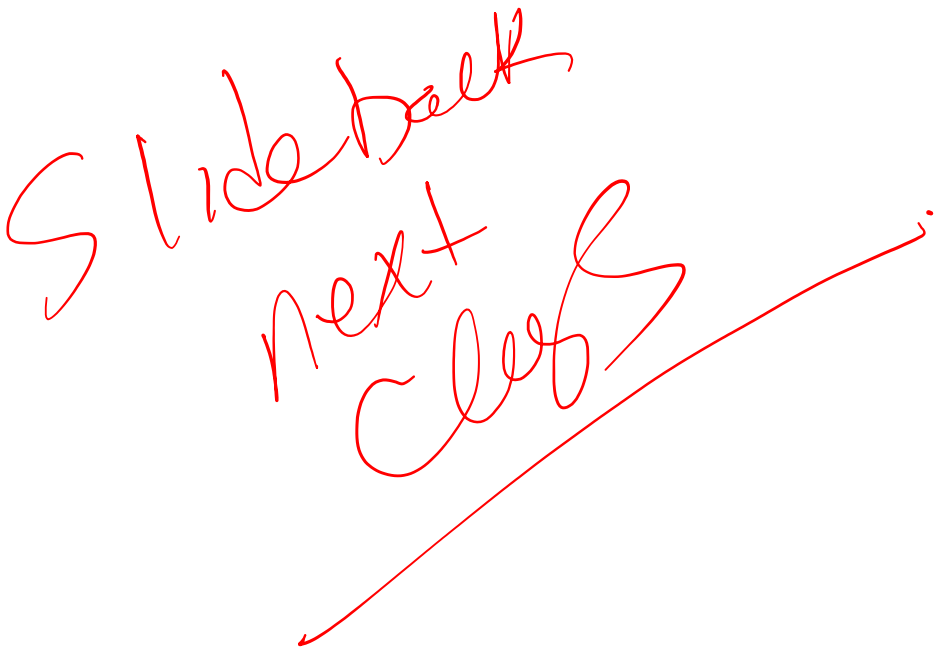
27

Book we will follow

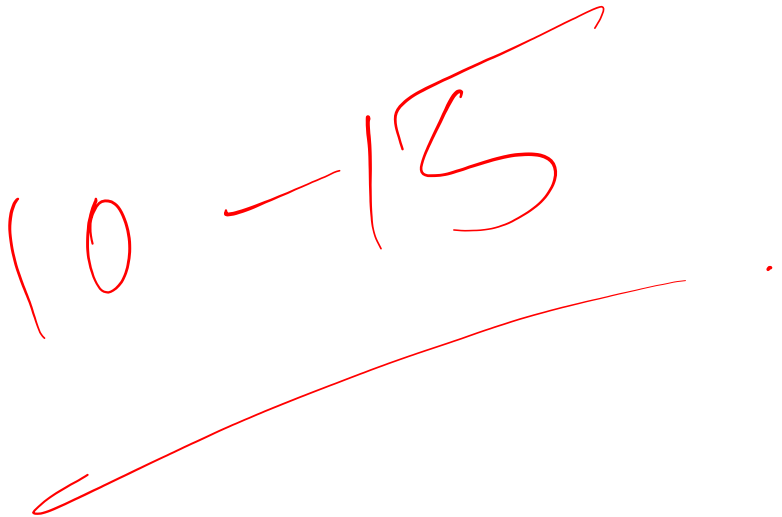


Any questions / clarifications?

What do you want to know by end of Sem?



How many already use MySQL? Oracle?



Basic Definitions

Database:

A collection of related data.

Data:

Known facts that can be recorded and have an implicit meaning.

Mini-world:

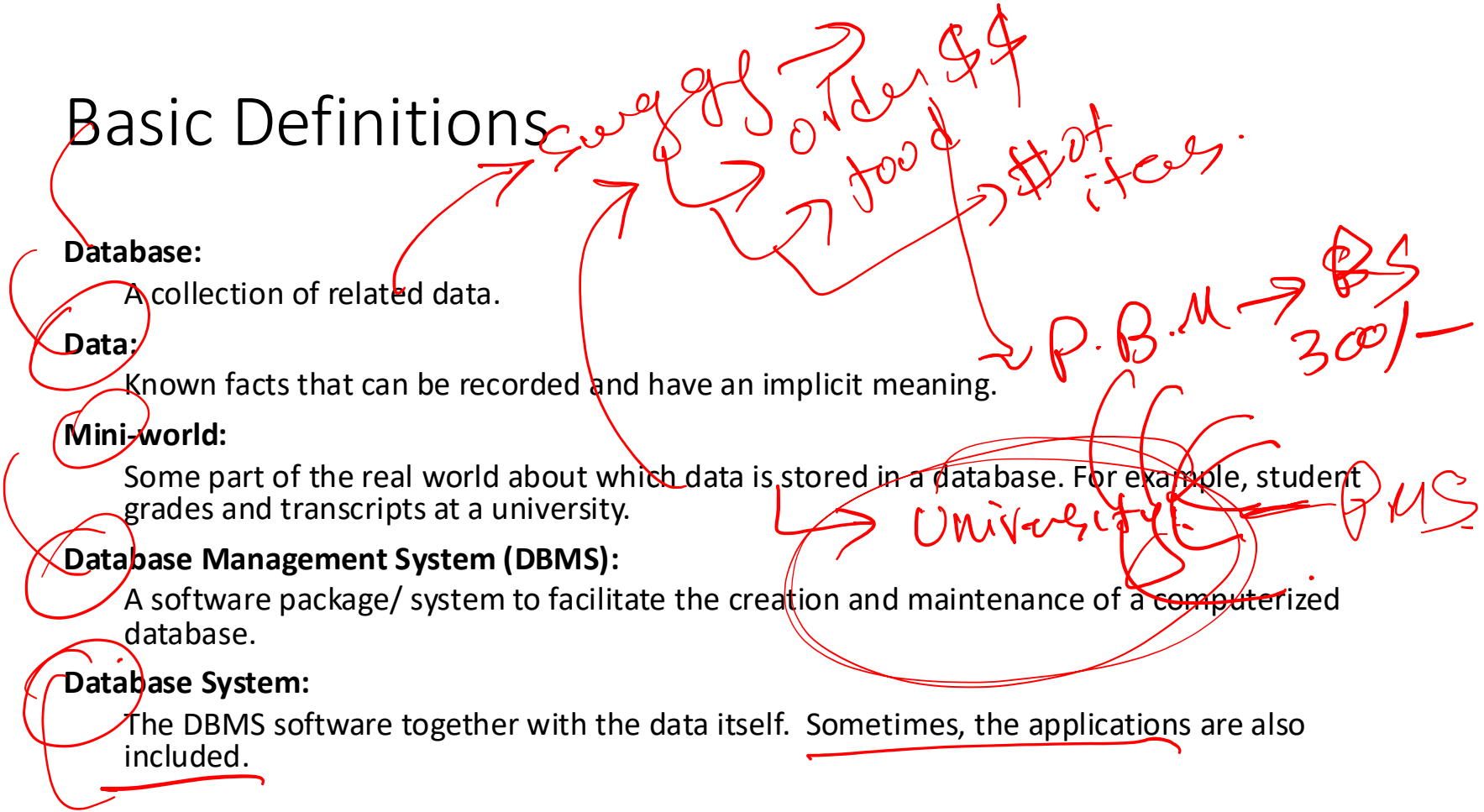
Some part of the real world about which data is stored in a database. For example, student grades and transcripts at a university.

Database Management System (DBMS):

A software package/ system to facilitate the creation and maintenance of a computerized database.

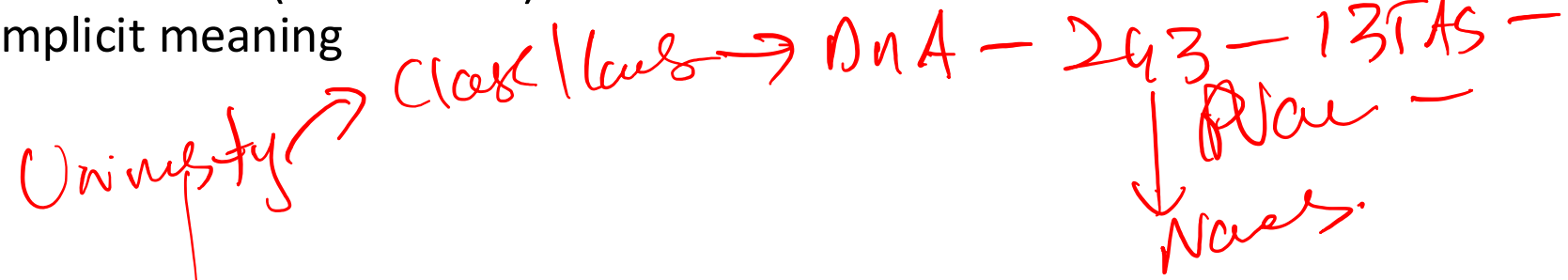
Database System:

The DBMS software together with the data itself. Sometimes, the applications are also included.

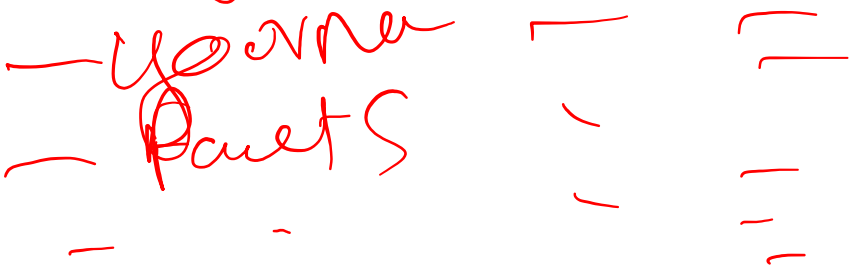


What is a Database?

Data: factual (undoubted) information that can be recorded and have implicit meaning



A database is a collection of related data



What is a Database?

A database has the following implicit properties:

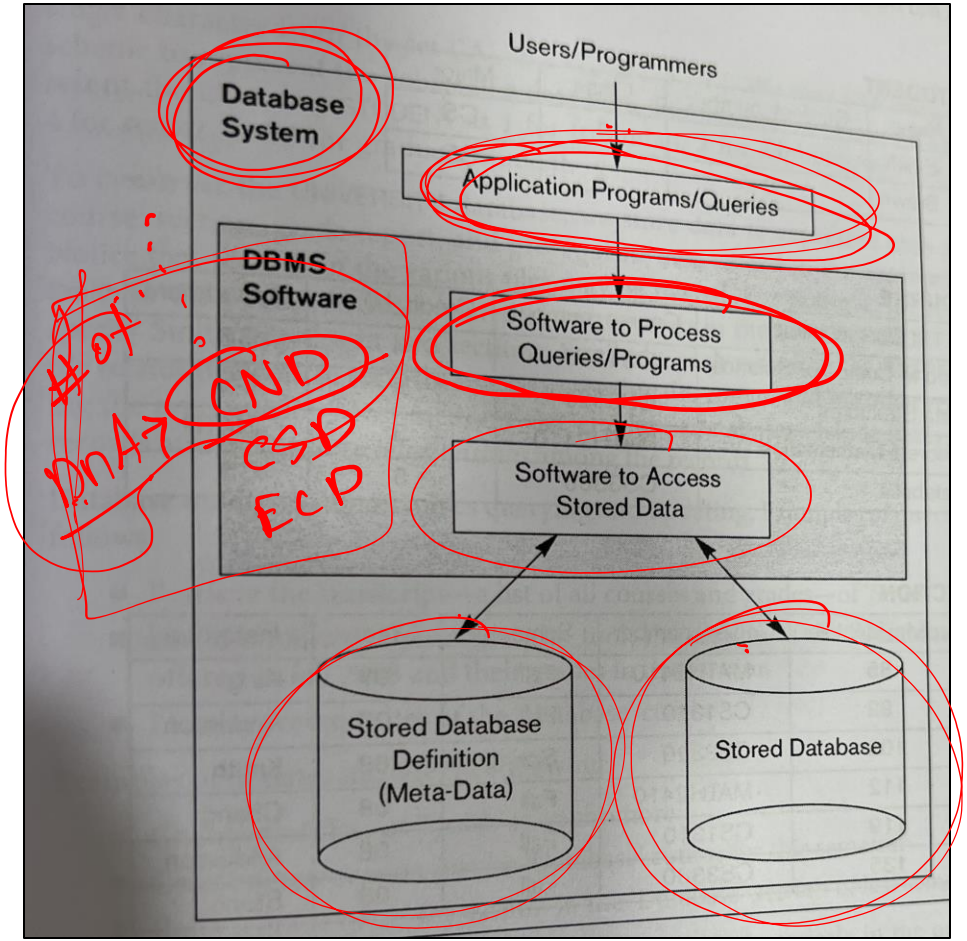
A database represents some aspect of the real world (mini-world or Universe of Discourse (UoD))

A database is a logically coherent (associated, related) collection of data with some inherent meaning

A database is designed, built and populated with data for a specific purpose

It has an intended group of users and some preconceived (already thought of) applications in which these users are interested

Simplified Database



Book

Example of Database: University

Data:

STUDENTs

COURSEs

SECTIONs (of COURSEs)

(academic) DEPARTMENTs

INSTRUCTORs

Relation:

SECTIONs are of specific COURSEs

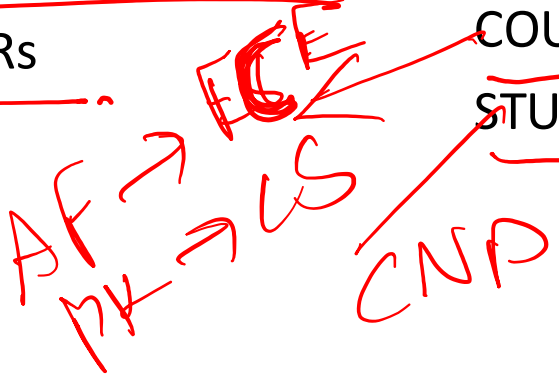
STUDENTs take SECTIONs

COURSEs have prerequisite COURSEs

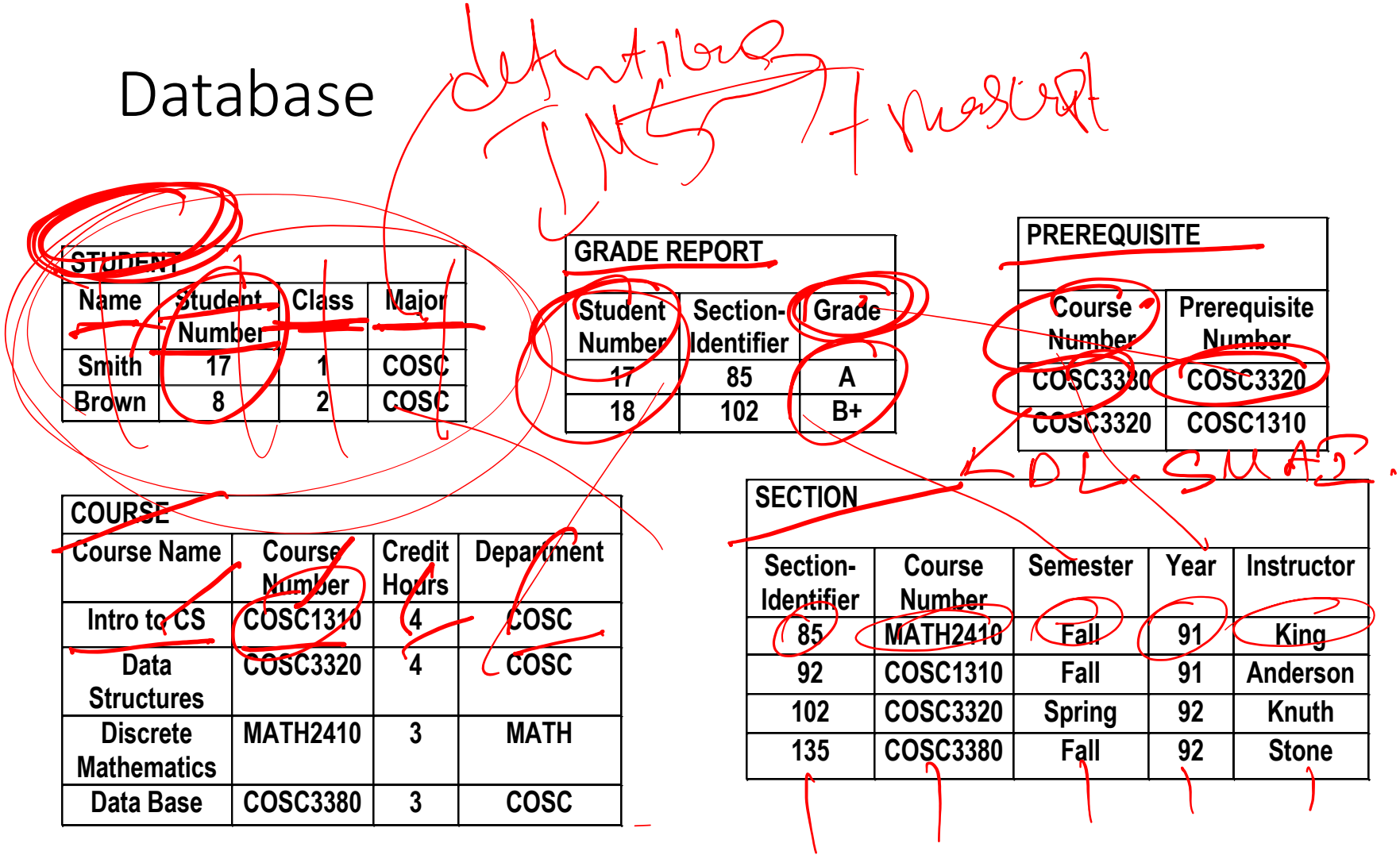
INSTRUCTORs teach SECTIONs

COURSEs are offered by DEPARTMENTs

STUDENTs major in DEPARTMENTs



Database



STUDENT

Name	Student Number	Class	Major
Smith	17	1	COSC
Brown	8	2	COSC

GRADE REPORT

Student Number	Section-Identifier	Grade
17	85	A
18	102	B+

PREREQUISITE

Course Number	Prerequisite Number
COSC3380	COSC3320
COSC3320	COSC1310

COURSE

Course Name	Course Number	Credit Hours	Department
Intro to CS	COSC1310	4	COSC
Data Structures	COSC3320	4	COSC
Discrete Mathematics	MATH2410	3	MATH
Data Base	COSC3380	3	COSC

SECTION

Section-Identifier	Course Number	Semester	Year	Instructor
85	MATH2410	Fall	91	King
92	COSC1310	Fall	91	Anderson
102	COSC3320	Spring	92	Knuth
135	COSC3380	Fall	92	Stone

Database catalogue

RELATIONS

Relation_name	No_of_columns
STUDENT	4
COURSE	4
SECTION	5
GRADE_REPORT	3
PREREQUISITE	2

COLUMNS

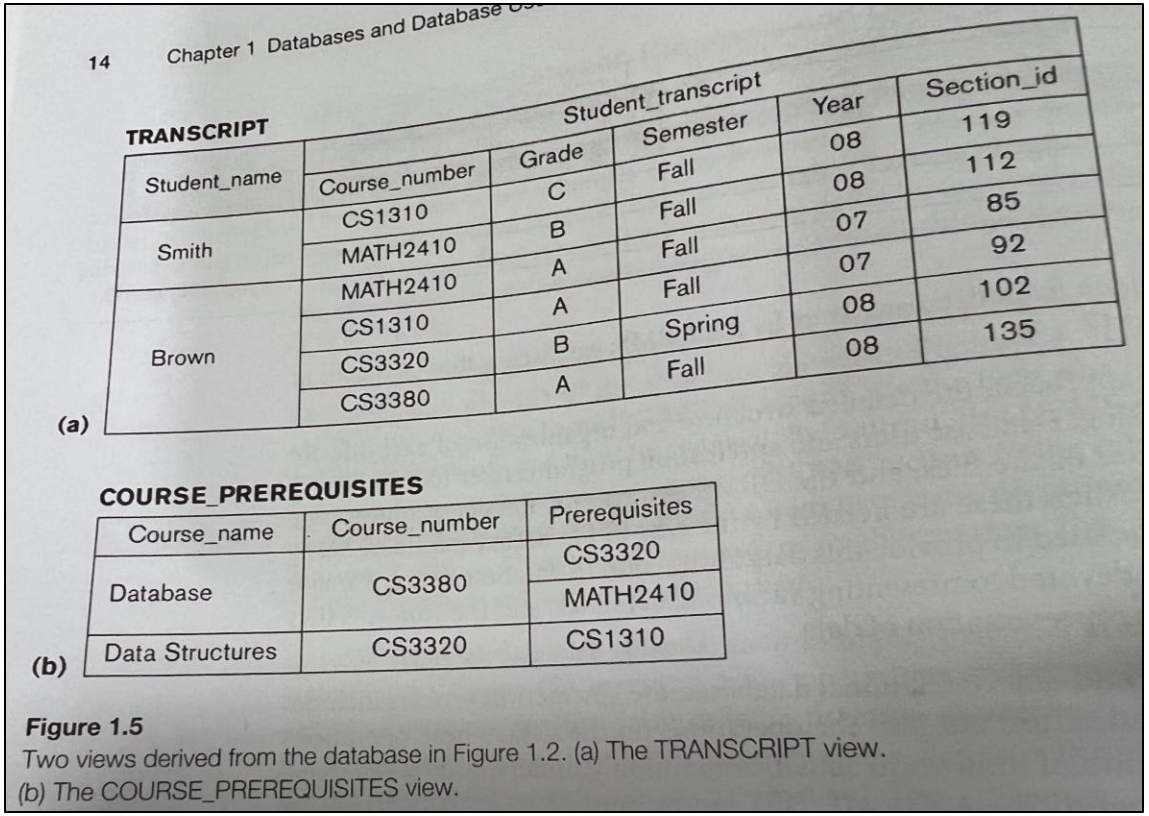
Column_name	Data_type	Belongs_to_relation
Name	Character (30)	STUDENT
Student_number	Character (4)	STUDENT
Class	Integer (1)	STUDENT
Major	Major_type	STUDENT
Course_name	Character (10)	COURSE
Course_number	XXXXNNNN	COURSE
....	....	.....
....	....	.....
....	....	.....
Prerequisite_number	XXXXNNNN	PREREQUISITE

Note: Major_type is defined as an enumerated type with all known majors.
XXXXNNNN is used to define a type with four alphabetic characters followed by fou

Book

Views

- Many users to DB
- Each users may require a different view
- View may be a subset or virtual data derived



TRANSCRIPT

Student_transcript					
Student_name	Course_number	Grade	Semester	Year	Section_id
Smith	CS1310	C	Fall	08	119
	MATH2410	B	Fall	08	112
Brown	MATH2410	A	Fall	07	85
	CS1310	A	Fall	07	92
	CS3320	B	Spring	08	102
	CS3380	A	Fall	08	135

(a)

COURSE_PREREQUISITES

Course_name	Course_number	Prerequisites
Database	CS3380	CS3320
		MATH2410
Data Structures	CS3320	CS1310

(b)

Figure 1.5
Two views derived from the database in Figure 1.2. (a) The TRANSCRIPT view. (b) The COURSE_PREREQUISITES view.

IMS DB.

D(A) By. Abed D(R) D. Y. K. Shome.

Online Transaction Processing (OLTP)

Multiuser DB

Concurrency control

Flight ticket booking, seats available

Transaction

Executing program or process that includes one or more database accesses, reading or updating of database records

Properties [ACID]

- Atomicity: either all are executed or none are executed [A/c A → A/c B]
- Consistency: any data written to a DB must be valid according to the defined rules [telephone number]
- Isolation: each transaction appears to execute in isolation, even though 100s may be executing at the same time [updating the seat preference]
- Durability: guarantees that once a transaction has been committed, it will remain committed even in the case of a system failure

Actors on the Scene: Day-to-Day use of DB

Database administrators

- authorizing access to DB, coordinating & monitoring its use, accountable for security breaches & response time

Database designers

- responsible for identifying the data to be stored in the DB, interact with potential group of users and develop *views* of the DB

End Users: Casual, naïve / parametric, sophisticated, stand-alone users

- Casual: occasional users, typically middle or high-level managers
- Naïve / parametric: constantly updating the db using *canned transaction,* done using mobile apps
  - bank tellers checking balances post withdrawals & deposits
  - reservation agents checking for availability
  - social media users post and read items on platforms

Any questions?

Bibliography / Acknowledgements

Instructor materials from Elmasri & Navathe 7e


pk.profgiri


Ponnurangam.kumaraguru


/in/ponguru


ponguru

Thank you for attending the class!!!


pk.guru@iiit.ac.in