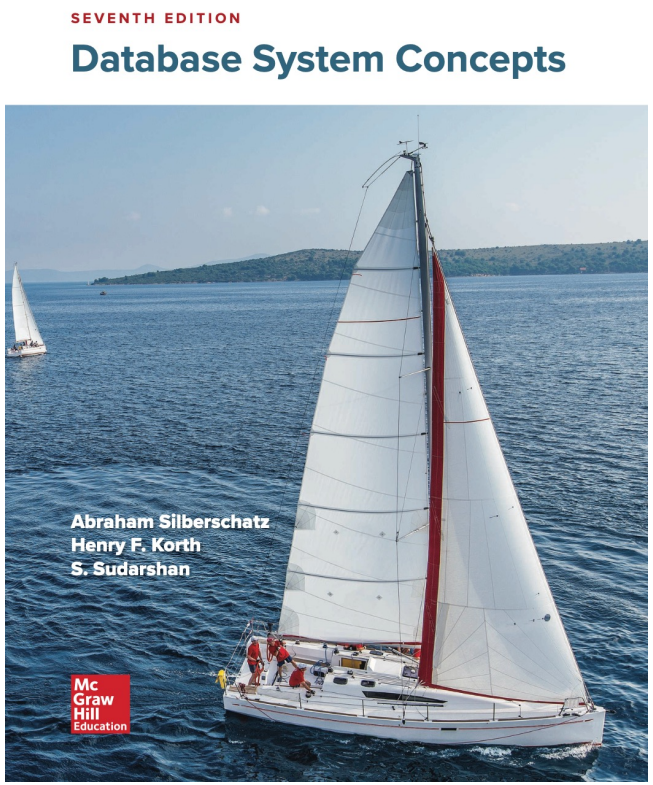
Lecture 1: Introduction

United International College

Course Information

- Text Book



Database System Concepts
Seventh Edition

Avi Silberschatz
Henry F. Korth
S. Sudarshan

McGraw-Hill Education
ISBN: 0073523321

Course Information

- Contact and Timetable

Section	Instructor	TA	Time		Venue
1001 1004	Dr. Zhiyuan Li goliathli@uic.edu.cn T3-502-R6	Mr. Shichen Zhang shichenzhang@uic.edu.cn T3-502-R26-H6	Lec	Mon 15:00-16:50	T5-505
			Lab	Wed 13:00-13:50	T8-303
			Tut	TBD	TBD
1002 1005	Dr. Jefferson Fong jeffersonfong@uic.edu.cn T3-602-R1	Mr. Steven Li mianli@uic.edu.cn T3-605-R25	Lec	Mon 10:00-11:50	T7-303
			Lab	Thu 15:00-15:50	T8-303
			Tut	TBD	TBD
1003 1006	Dr. Tianhui Meng tianhuimeng@uic.edu.cn T6-403-R7	Ms. Lily Lin chengyanlin@uic.edu.cn T3-602-R25	Lec	Wed 9:00-10:50	T7-303
			Lab	Fri 9:00-9:50	T29-101
			Tut	TBD	TBD

Course assessment

- 3 Assignments: 10%
- Midterm: 20%
- Lab: 10%
- Group project: 20%
- Final Exam: 40%

Course assessment

Regulations

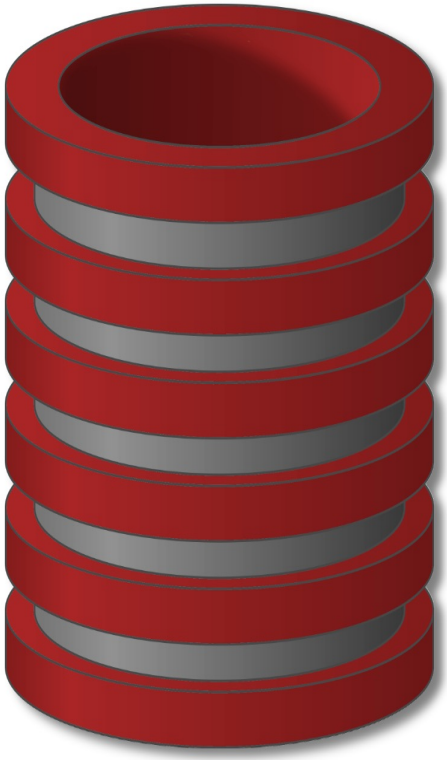
- MR sections and FE sections will be graded separately.
- Low final exam score will downgrade the letter grade.

Final Exam Score	Highest Letter Grade
0 - 19	F
20 - 24	D
25 - 29	C-

Lecture Weekly Schedule

Week	Lecture	Readings	Lab
1	1. Introduction	Chap. 1	1. XAMPP
2	2. Basic ER model	Chap. 6.1-3, 6.5	2. Basic Query
3	3. ER Constraints	Chap. 6.4	3. Cross Table Queries
4	4. Extended ER	Chap. 6.5.3, 6.8	4. Join
5	5. ER Design Issues	Chap. 6.9	5. Aggregation
6	6. Relational Model	Chap. 2.1-4, 6.8.6	6. Data Definition
7	7. Exercises	-	7. Data Modification
8	Reading Week		
9	8. Design Purpose	Chap. 7.1	8. Null Values & Subquery 1
10	9. Functional Dependencies	Chap. 7.4	9. Set Operations & Subquery 2
11	10. BCNF	Chap. 7.3, 7.5	10. Correlation
12	11. 3NF	Chap. 7.3, 7.5	11. Foreign key & Triggers
13	12. MVDs (Optional)	Chap. 7.6	12. Constraints
14	13. XML Database	Chap. 30	13. Triggers
15	Project Presentation		

Introduction to Databases



Whether you know it or not, you're using a database every day

Database Management System (DBMS)

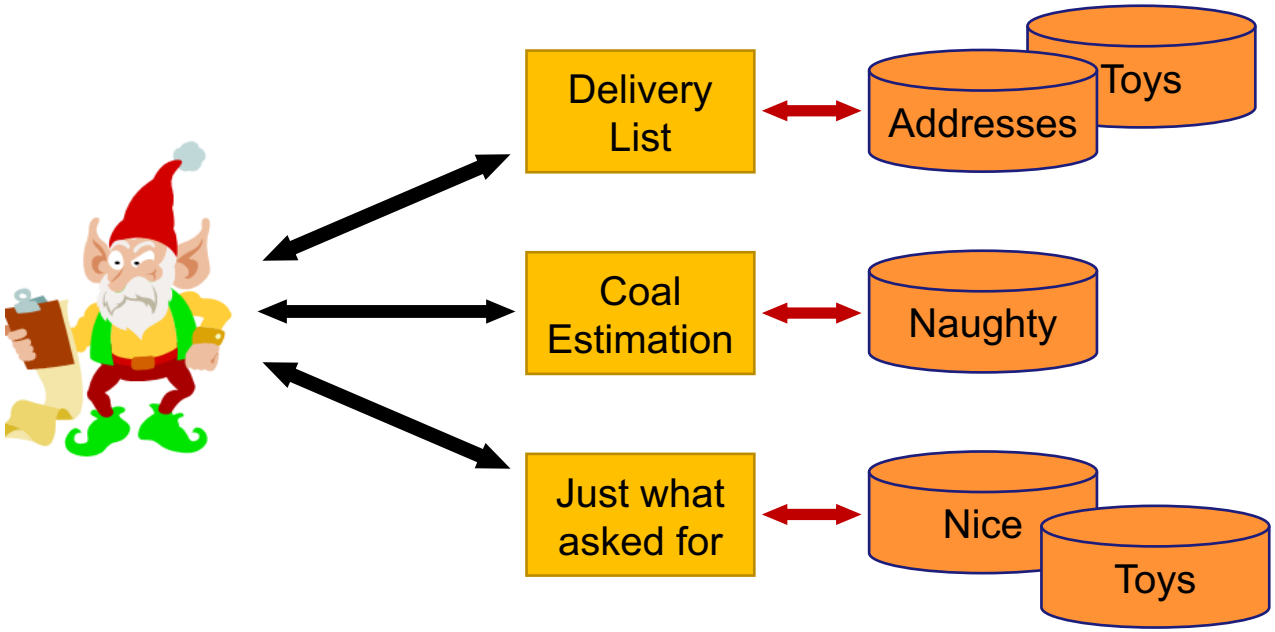
- Database Applications
  - Banking: transactions
  - Airlines: reservations, schedules
  - Universities: registration, grades
  - Sales: customers, products, purchases
  - Online retailers: order tracking, customized recommendations
  - Manufacturing: production, inventory, orders, supply chain
  - Human resources: employee records, salaries, tax deductions
  - Social media sites: posts, friendship relationship, activities
- Database can be very large.
- Database touch all aspects of our lives.

University Database Example

- Course Management
  - Add students, instructors, and courses
  - Register students for courses, and generate class rosters
  - Assign grades to students, compute grade point averages (GPA) and generate transcripts
- Provide basic features necessary for data access
  - Shared access by a community of uses
  - Well-defined schema for data access
  - Support query language

File systems

- In the early days, applications were built directly on top of file systems.



Drawbacks of file systems

- Data redundancy and inconsistency
  - Multiple file formats, duplication of information in different files.
- Difficulty in accessing data
  - Need to write a new program to carry out each new task.
- Data isolation
  - Multiple files and formats
- Integrity problems
  - Integrity constraints (e.g., account balance ≥ 0) become buried in program code rather than being stated explicitly.
  - Hard to add new constraints or change existing ones.

Drawbacks of file systems

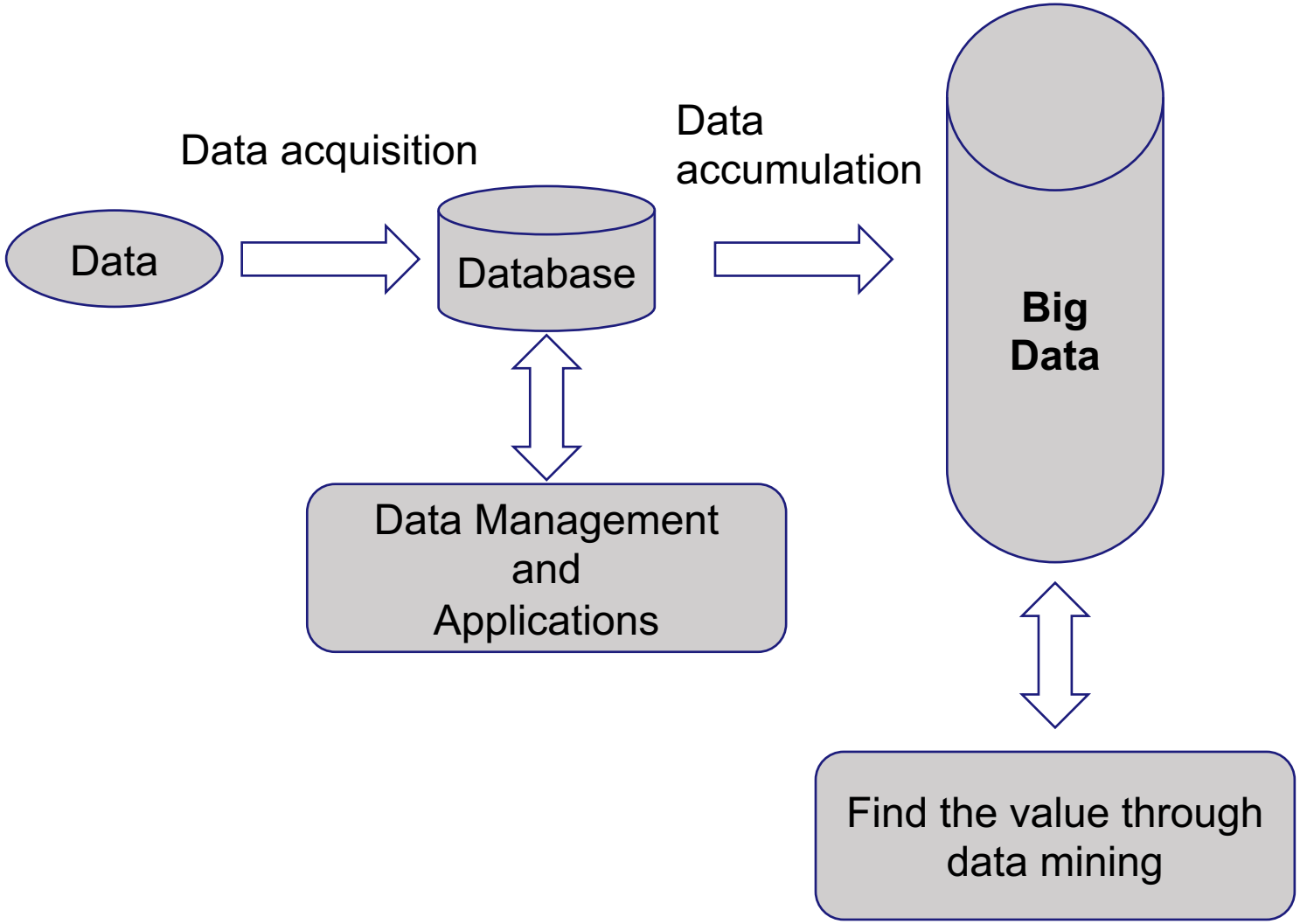
- Atomicity of updates
  - Failures may leave database in an inconsistent state with partial updates carried out.
  - Example: Transfer of funds from one account to another should either complete or not happen at all.
- Concurrent access by multiple users
  - Concurrent access needed for performance
  - Uncontrolled concurrent accesses can lead to inconsistencies
    - Example: Two students want to choose the same course that has only one vacancy left.
- Security problems
  - Hard to provide user access to some, but not all, data

Database system offer solutions to all the above problems.

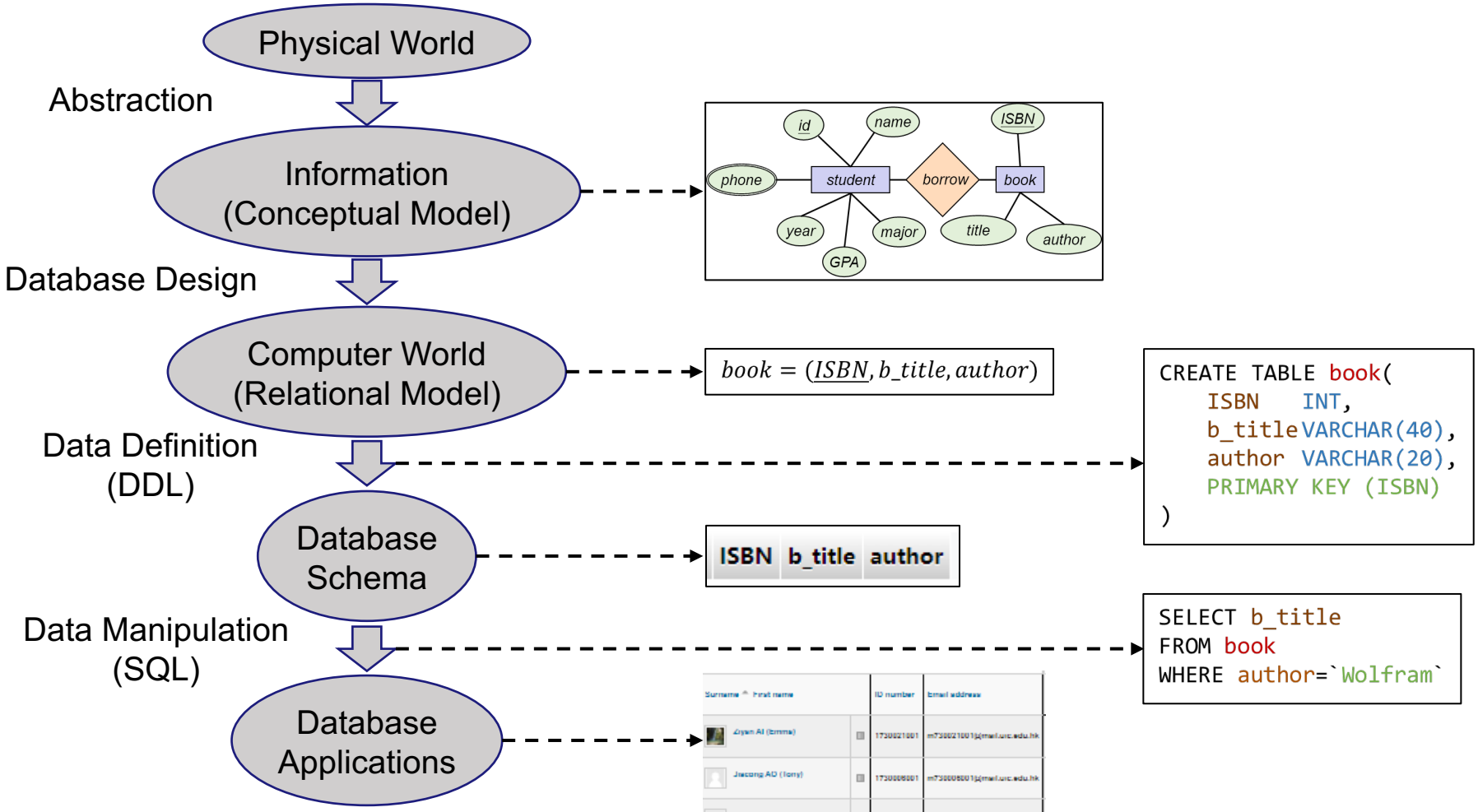
DBMS provides...

- Efficient
- Reliable
- Convenient
- Safe
- Multi-user (Storage of and access to)
- Massive amounts of persistent data

In general

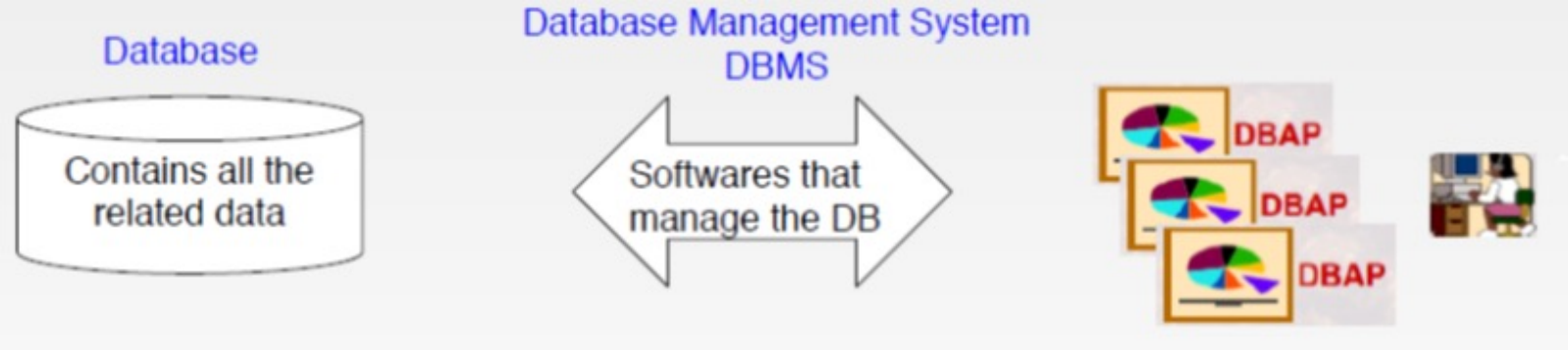

Data acquisition
Data accumulation
Data
Database
Big Data
Data Management and Applications
Find the value through data mining

Key Concepts



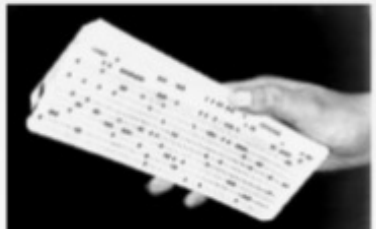
Key Concepts

- Data modeling
- Schema vs. Data instances
- Data definition language (DDL)
- Data manipulation or query language (DML)
- DBMS implementer
- Database designer
- Database application developer
- Database administrator



History of Database Systems

- 1950s and early 1960s:
  - Data processing using magnetic tapes for storage
    - Tapes provide only sequential access
  - Punched cards for input
- Late 1960s and 1970s:
  - Hard disks allow direct access to data
  - Network and hierarchical data models in widespread use
  - Ted Codd defines the relational data model
    - Would win the ACM Turing Award for this work
    - IBM Research begins System R prototype
    - UC Berkeley begins Ingres prototype

History (cont.)

- 1980s:
  - Research relational prototypes evolve into commercial systems
    - SQL becomes industrial standard
  - Parallel and distributed database systems
  - Object-oriented database systems
- 1990s:
  - Large decision support and data-mining applications
  - Large multi-terabyte data warehouses
  - Emergence of Web commerce
- 2000s:
  - XML and XQuery standards
  - Automated database administration

Database Popularity Rank 2017 - 2018

Rank Feb 2018	Rank Jan 2018	Rank Feb 2017	DBMS	Database Model	Score Feb 2018	Score Jan 2018	Score Feb 2017
1.	1.	1.	Oracle	Relational DBMS	1303.28	-38.66	-100.55
2.	2.	2.	MySQL	Relational DBMS	1252.47	-47.24	-127.83
3.	3.	3.	Microsoft SQL Server	Relational DBMS	1122.04	-26.03	-81.42
4.	4.	4.	PostgreSQL	Relational DBMS	388.38	+2.19	+34.70
5.	5.	5.	MongoDB	Document store	336.42	+5.47	+0.92
6.	6.	6.	DB2	Relational DBMS	189.97	-0.30	+2.07
7.	7.	↑ 8.	Microsoft Access	Relational DBMS	130.07	+3.37	-3.32
8.	↑ 9.	↑ 10.	Redis	Key-value store	127.02	+3.88	+12.98
9.	↑ 10.	↑ 11.	Elasticsearch	Search engine	125.32	+2.76	+17.01
10.	↓ 8.	↓ 7.	Cassandra	Wide column store	122.78	-1.10	-11.60
11.	11.	↓ 9.	SQLite	Relational DBMS	117.27	+3.02	+1.96
12.	12.	12.	Teradata	Relational DBMS	72.99	+0.36	-2.60
13.	↑ 15.	↑ 16.	Splunk	Search engine	67.27	+3.27	+11.24
14.	14.	14.	Solr	Search engine	63.87	-0.50	-3.81
15.	↓ 13.	↓ 13.	SAP Adaptive Server	Relational DBMS	63.49	-1.98	-8.25
16.	16.	↓ 15.	HBase	Wide column store	61.70	+0.07	+2.46

Database Popularity Rank 2019 - 2020

Rank Jan 2020	Rank Dec 2019	Rank Jan 2019	DBMS	Database Model	Score Jan 2020	Score Dec 2019	Score Jan 2019
1.	1.	1.	Oracle	Relational, Multi-model	1346.68	+0.29	+77.85
2.	2.	2.	MySQL	Relational, Multi-model	1274.65	-1.01	+120.39
3.	3.	3.	Microsoft SQL Server	Relational, Multi-model	1098.55	+2.35	+58.29
4.	4.	4.	PostgreSQL	Relational, Multi-model	507.19	+3.82	+41.08
5.	5.	5.	MongoDB	Document, Multi-model	426.97	+5.85	+39.78
6.	6.	6.	IBM Db2	Relational, Multi-model	168.70	-2.65	-11.15
7.	7.	↑ 8.	Elasticsearch	Search engine, Multi-model	151.44	+1.19	+8.00
8.	8.	↓ 7.	Redis	Key-value, Multi-model	148.75	+2.51	-0.27
9.	9.	9.	Microsoft Access	Relational	128.58	-0.89	-13.04
10.	↑ 11.	10.	SQLite	Relational	122.14	+1.78	-4.66
11.	↓ 10.	11.	Cassandra	Wide column	120.66	-0.04	-2.32
12.	12.	12.	Splunk	Search engine	88.67	-1.85	+7.25
13.	13.	13.	MariaDB	Relational, Multi-model	87.45	+0.66	+8.63
14.	14.	↑ 15.	Hive	Relational	84.24	-1.81	+14.33
15.	15.	↓ 14.	Teradata	Relational, Multi-model	78.29	-0.21	+2.10
16.	16.	↑ 20.	Amazon DynamoDB	Multi-model	62.02	+0.39	+6.93

Database Popularity Rank 2020 - 2021

Rank Sep 2021	Rank Aug 2021	Rank Sep 2020	DBMS	Database Model	Score Sep 2021	Score Aug 2021	Score Sep 2020
1.	1.	1.	Oracle	Relational, Multi-model	1271.55	+2.29	-97.82
2.	2.	2.	MySQL	Relational, Multi-model	1212.52	-25.69	-51.72
3.	3.	3.	Microsoft SQL Server	Relational, Multi-model	970.85	-2.50	-91.91
4.	4.	4.	PostgreSQL	Relational, Multi-model	577.50	+0.45	+35.22
5.	5.	5.	MongoDB	Document, Multi-model	496.50	-0.04	+50.02
6.	6.	↑ 7.	Redis	Key-value, Multi-model	171.94	+2.05	+20.08
7.	7.	↓ 6.	IBM Db2	Relational, Multi-model	166.56	+1.09	+5.32
8.	8.	8.	Elasticsearch	Search engine, Multi-model	160.24	+3.16	+9.74
9.	9.	9.	SQLite	Relational	128.65	-1.16	+1.98
10.	↑ 11.	10.	Cassandra	Wide column	118.99	+5.33	-0.18
11.	↓ 10.	11.	Microsoft Access	Relational	116.94	+2.10	-1.51
12.	12.	12.	MariaDB	Relational, Multi-model	100.70	+1.72	+9.09
13.	13.	13.	Splunk	Search engine	91.61	+1.01	+3.71
14.	14.	↑ 15.	Hive	Relational	85.58	+1.64	+14.41
15.	15.	↑ 17.	Microsoft Azure SQL Database	Relational, Multi-model	78.26	+3.11	+17.81
16.	16.	16.	Amazon DynamoDB	Multi-model	76.93	+2.03	+10.75

End of Chapter 1