

Indice generale

Ringraziamenti	xii
Prefazione	xiii
A chi è rivolto questo manuale?	xiii
Capitolo 1 Le basi di dati	1
Che cos'è un database?.....	2
Breve storia dei database	3
A che cosa servono i database?	9
L'importanza di gestire (bene) i dati.....	10
Sistemi centralizzati	10
Gestione strutturata dei dati	10
Conclusioni	13
Che cosa abbiamo imparato	14
Capitolo 2 Nozioni di base.....	15
Dati.....	16
Database	16
DBMS.....	17
Database relazionale.....	18
Database gerarchico	19
Database a grafo.....	20
Database a oggetti.....	20
Modelli di gestione dei dati.....	22
Il modello ER	24
Entità	24
Attributo	25
Relazione	27
Che cosa abbiamo imparato	33

Capitolo 3	SQL: le basi	35
	SQL come linguaggio di base.....	35
	Strumenti di base	36
	Istruzioni	37
	SELECT	37
	WHERE	39
	AND	40
	OR	41
	NOT	42
	ORDER BY	43
	Operatori	44
	BETWEEN.....	45
	IN.....	47
	TOP.....	48
	DISTINCT	49
	COUNT()	50
	SUM().....	51
	AVG().....	51
	MIN() e MAX()	51
	GROUP BY	52
	HAVING	54
	ANY,ALL	55
	LIKE	57
	Wildcard	58
	AS.....	59
	UNION.....	60
	INTERSECT	63
	EXISTS.....	64
	Gestire un database	65
	CREATE	65
	ALTER	67
	INSERT.....	67
	UPDATE	69
	DELETE.....	70
	DROP	71
	Chiavi.....	72
	Tipi di dato	75
	Istruzioni: JOIN.....	77
	Tipologie	79
	Esempio	81
	Che cosa abbiamo imparato	84
Capitolo 4	SQL: avanzato.....	85
	Viste	85
	Stored procedure	87

Indici.....	89
Transazioni	92
Atomicità	93
Coerenza	93
Isolamento.....	94
Durezza.....	94
Livelli di isolamento.....	96
Read uncommitted.....	97
Read committed.....	97
Lettura ripetibile.....	97
Serializzabile.....	98
Trigger	98
Che cosa abbiamo imparato	100

Capitolo 5 MySQL.....101

Installazione per Windows.....	102
Installazione per Ubuntu 18.x e 20.x.....	107
Installazione per macOS.....	111
Utilizzo pratico.....	117
Caso di studio: l'università.....	117
Svolgimento	118
Diagramma ER	118
Creazione del database.....	122
Creazione delle tabelle.....	124
Calcolo della media dei voti ottenuti in un certo periodo per un determinato corso	128
Calcolo del numero di studenti che non hanno superato l'esame, con periodo di riferimento del corso e professore	130
Verifica della presenza di eccellenze nel dipartimento, ossia di studenti con una media superiore al 28 e che abbiano conseguito almeno tre voti pari a 30 e lode	132
Numero di crediti di un dato studente, fornendo la matricola	134
Calcolo dinamico del numero di iscritti per ogni corso	135
Conclusioni	136
Che cosa abbiamo imparato	136

Capitolo 6 PostgreSQL.....137

Installazione per Windows.....	138
Installazione per Ubuntu 18.x e 20.x.....	141
Installazione per macOS.....	142
Utilizzo pratico.....	146
Caso di studio: catena di supermercati.....	146
Svolgimento	147
Diagramma ER	147

Creazione del database	151
Calcolo della media dei guadagni in un certo periodo per un determinato negozio	155
Calcolo dei clienti che effettuano una spesa inferiore a 200 euro complessive nell'ultimo mese	157
Verifica dell'esistenza di un prodotto	160
Che cosa abbiamo imparato	161

Capitolo 7 Database NoSQL163

Che cos'è un database NoSQL.....	163
Database chiave-valore	164
Orientato al documento	165
A grafo	165
Orientato alle colonne	167
Perché scegliere un database NoSQL.....	168
Quando scegliere un database NoSQL	168
Che cosa abbiamo imparato	169

Capitolo 8 MongoDB.....171

Installazione per Windows.....	173
Installazione per Ubuntu.....	175
Installazione per macOS.....	177
Primi passi	177
Creazione di un database.....	180
Creazione di una collezione.....	180
Cercare tra i documenti	184
Cercare determinati campi	191
Ordinamenti.....	193
Eliminare documenti	194
Aggregazione di dati.....	195
Gestire le relazioni	198
Denormalizzazione	201
Quando usare MongoDB.....	202
Schema-less	202
Durabilità	203
Transazioni	204
Geospazialità.....	204
Prestazioni e utility	204
Backup e restore	205
Conclusioni	207
Che cosa abbiamo imparato	208

Capitolo 9	Elasticsearch.....	209
	Che cos'è.....	209
	Nodo	210
	Cluster	210
	Indice	211
	Documento	211
	Shard.....	211
	Replica.....	211
	Installazione su Ubuntu 18.x e 20.x	212
	Installazione su Windows	213
	Usare le API	217
	Operazioni CRUD	219
	Inserire un documento	220
	Recuperare un documento	222
	Filtrare la ricerca.....	226
	Modificare un documento	234
	Aggregazioni	236
	Cancellare un documento.....	238
	Cancellare un indice	240
	Gestione del cluster.....	240
	Modellare i dati	243
	Join lato applicazione.....	244
	Denormalizzazione	246
	Oggetti annidati.....	248
	Relazioni genitore-figlio.....	249
	Che cosa abbiamo imparato	252
Capitolo 10	La progettazione.....	253
	Perché la progettazione è importante?	253
	Strategie	255
	Metodo top-down.....	255
	Metodo bottom-up	256
	Best practice	259
	Prendi in considerazione la prospettiva di tutti	259
	Il nome è importante.....	259
	La normalizzazione è la chiave	260
	Scrivere documentazione.....	260
	La privacy al primo posto	260
	Pensa alle necessità a lungo termine	260
	Codice e utilizzo di stored procedure.....	261
	Dedica tempo alla modellazione e progettazione di database	261
	Metti alla prova il design	261
	Scegli un database che “calzi a pennello”.....	261
	Che cosa abbiamo imparato	264

Capitolo 11 Big data.....265

Come nascono i big data.....	265
Big data = big impact	272
La sfida	272
Rapida crescita	272
Conservazione.....	273
Sincronizzazione tra diverse sorgenti	273
Sicurezza	273
Inaffidabilità dei dati	273
Processo di elaborazione	273
Big data stack.....	274
Apache Flume	275
Sqoop.....	276
Apache Kafka	276
HDFS	277
Apache Hbase.....	278
Framework	278
Hadoop.....	278
Spark.....	280
Flink	281
Storm.....	282
Samza.....	282
Applicazioni	283
Istituzioni governative.....	283
Assistenza sanitaria.....	284
Media.....	284
Assicurazioni	284
Che cosa abbiamo imparato	285

Capitolo 12 Data on cloud287

DynamoDB.....	290
Chiave primaria.....	290
Indici secondari	291
Utilizzo pratico.....	291
Usare le API	295
Operazioni CRUD.....	299
Eseguire query.....	302
S3.....	303
Hosting di un sito web.....	304
Amazon Keyspaces.....	311
Cassandra Data Model	313
CQL	314
Che cosa abbiamo imparato	320

Capitolo 13 Data Quality Management.....	321
Come funziona.....	322
Qual è la definizione di qualità dei dati?.....	322
I pilastri della gestione della qualità dei dati.....	323
Le persone.....	323
Profilazione dei dati.....	323
Definizione della qualità dei dati.....	324
Reportistica dei dati.....	324
Correzione dei dati.....	324
Metriche di data quality.....	325
Accuratezza.....	325
Coerenza.....	325
Completezza.....	325
Integrità.....	325
Tempestività.....	326
I vantaggi.....	326
Processo di data quality.....	327
Data integrity.....	327
Data cleansing.....	328
Data matching.....	330
Data enrichment.....	333
Data validation.....	334
Che cosa abbiamo imparato.....	335
Capitolo 14 Conclusioni.....	337
Aspetti futuri.....	340
Indice analitico.....	343