Algorithm for Developing ER Model

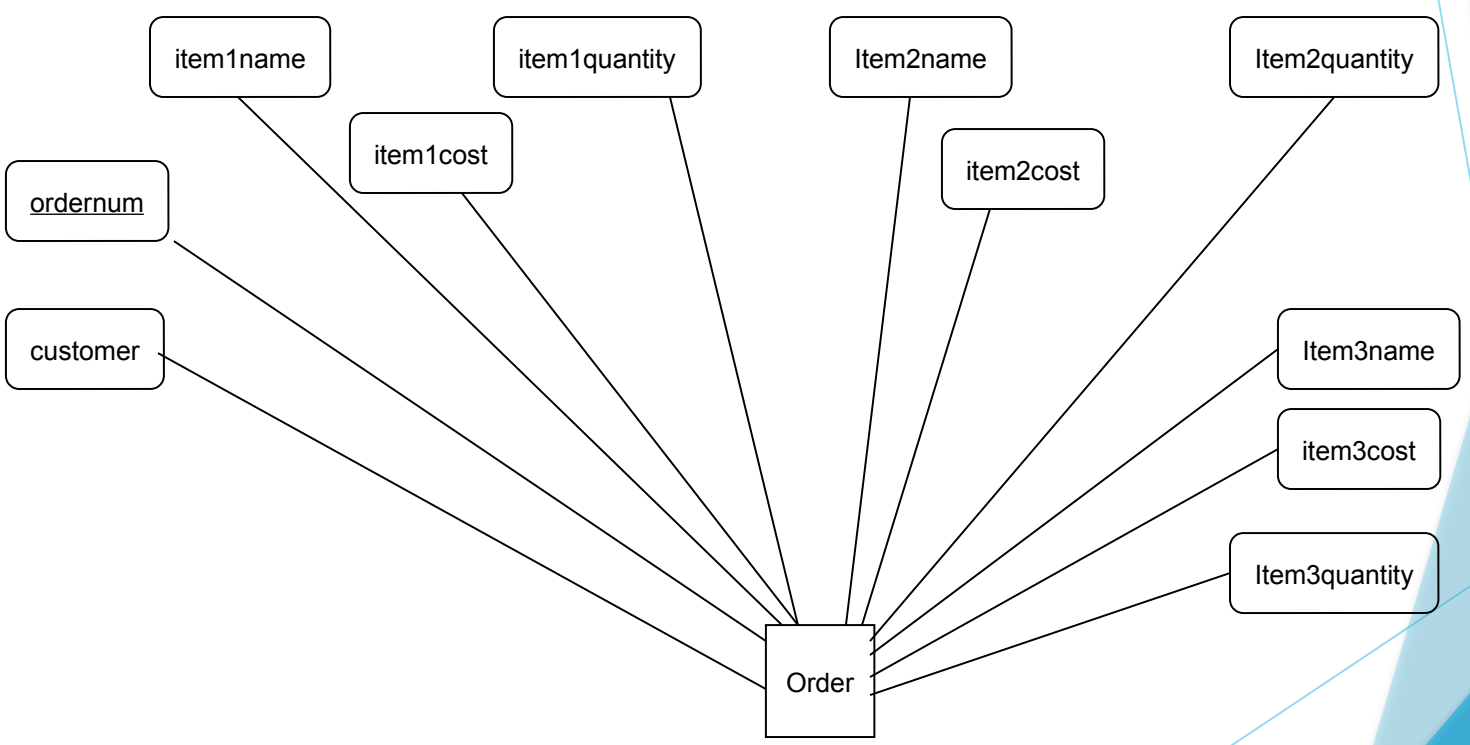
1. Create detailed narrative of organization's order of operation
2. Identify business rules based on operations
3. Identify entities, attributes and relationships
4. Develop ERD
5. Identify primary keys and foreign keys for each entity (if applicable)
6. REVISE as necessary

Normalization

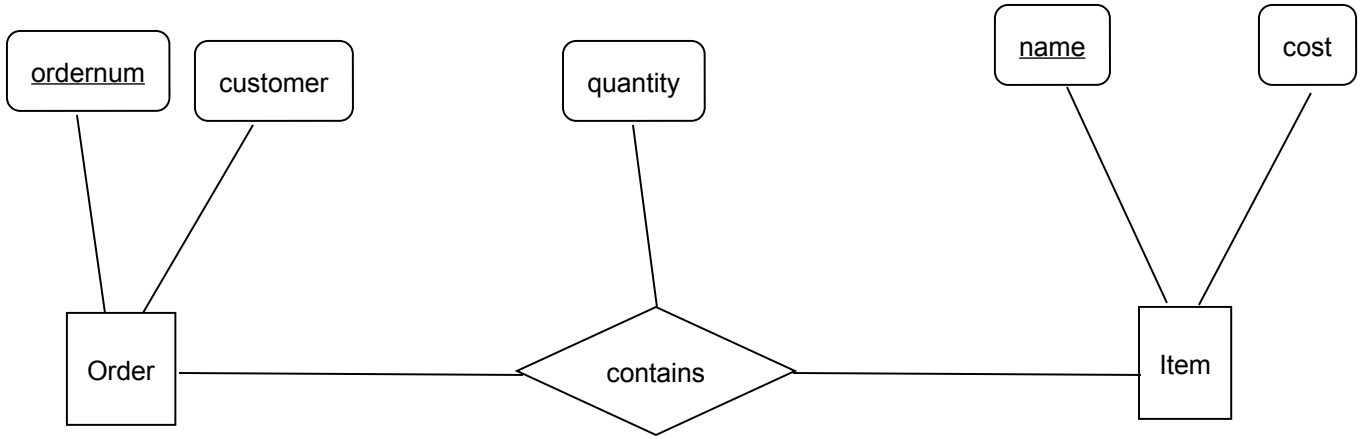
Normalization

- A technique used in designing relational databases
  - Removes redundancies
  - Minimizes dependencies
  - Removes anomalies
    - Insert
    - Update
    - Delete
- 5 Normal Forms
  - 3 are most important

Before Normalization



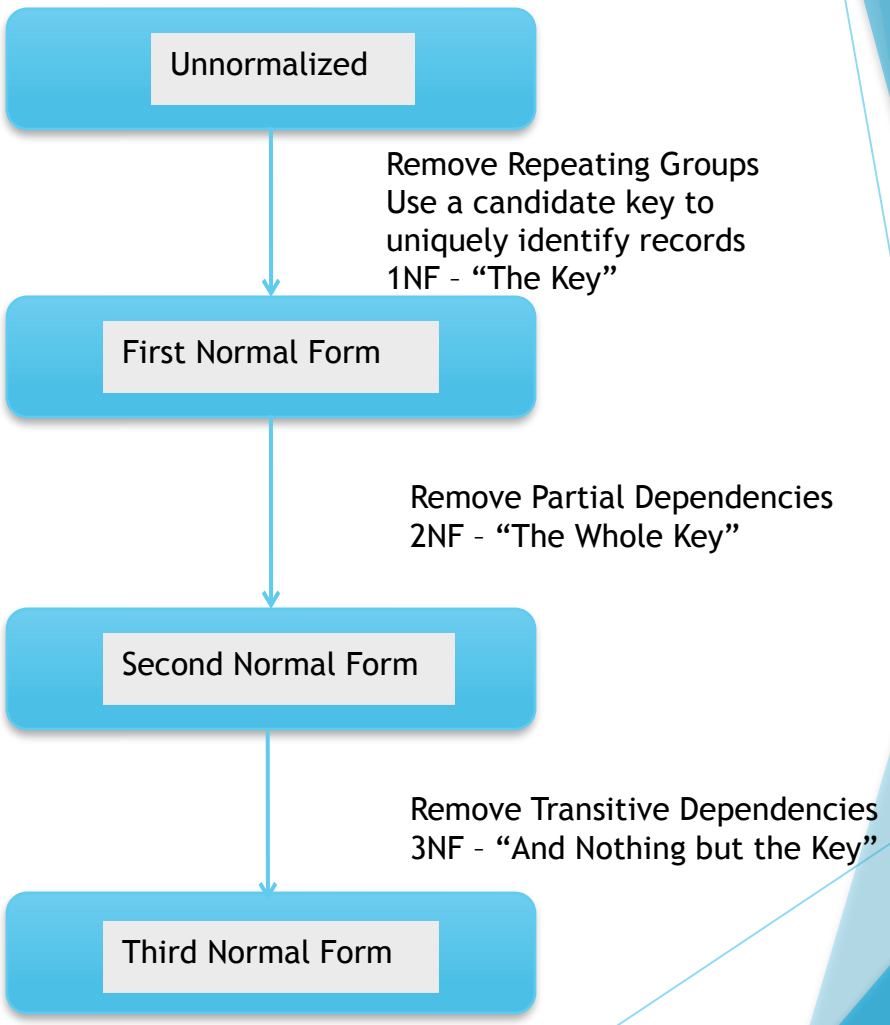
After Normalization



Anomalies

Supplier_Num	City	Part_Num	Qty
S1	London	P1	100
S1	London	P2	100
S2	Paris	P1	200
S2	Paris	P2	200
S3	Paris	P2	300
S4	London	P2	400
S4	London	P4	400
S4	London	P5	400

Normalization



Unnormalized

PROJ_NUM	PROJ_NAME	EMP_NUM	EMP_NAME	JOB_CLASS	CHG_HOUR	HOURS
15	Evergreen	103	June E. Arbough	Elect. Engineer	84.50	23.80
		101	John G. News	Database Designer	105.00	19.40
		105	Alice K. Johnson *	Database Designer	105.00	35.70
		106	William Smithfield	Programmer	35.75	12.60
		102	David H. Senior	Systems Analyst	96.75	23.80
18	Amber Wave	114	Annelise Jones	Applications Designer	48.10	24.60
		118	James J. Frommer	General Support	18.36	45.30
		104	Anne K. Ramoras *	Systems Analyst	96.75	32.40
		112	Darlene M. Smithson	DSS Analyst	45.95	44.00
22	Rolling Tide	105	Alice K. Johnson	Database Designer	105.00	64.70
		104	Anne K. Ramoras	Systems Analyst	96.75	48.40
		113	Delbert K. Joenbrood *	Applications Designer	48.10	23.60
		111	Geoff B. Wabash	Clerical Support	26.87	22.00
		106	William Smithfield	Programmer	35.75	12.80
25	Starflight	107	Maria D. Alonzo	Programmer	35.75	24.60
		115	Travis B. Bawangi	Systems Analyst	96.75	45.80
		101	John G. News *	Database Designer	105.00	56.30
		114	Annelise Jones	Applications Designer	48.10	33.10
		108	Ralph B. Washington	Systems Analyst	96.75	23.60
		118	James J. Frommer	General Support	18.36	30.50
		112	Darlene M. Smithson	DSS Analyst	45.95	41.40

Unnormalized

- Unnormalized
  - **Project(Project_Num, Project_Name, (Emp_Num, Emp_Name), (Job_Class, Chg_Hour))**

1NF

- 1NF – *The Key*
  - Let Project_Num and Emp_Num be the composite key
  - **Project(Project_Num, Project_Name, Emp_Num, Emp_Name, Job_Class, Chg_Hour)**

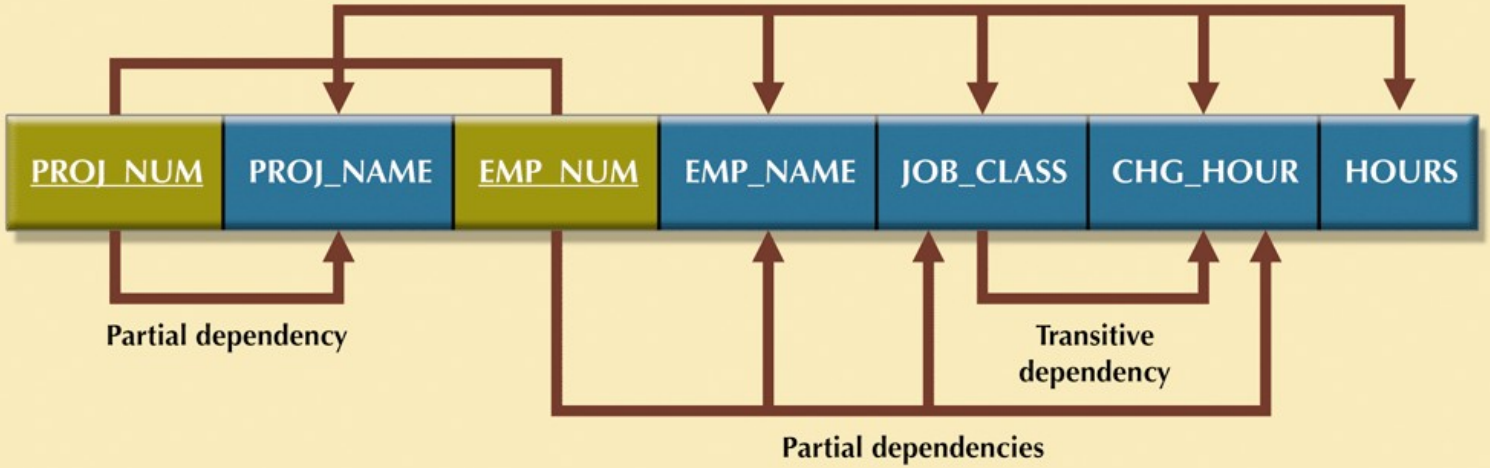
FIGURE 6.2 A table in first normal form

Table name: DATA_ORG_1NF **Database name: Ch06_ConstructCo**

PROJ_NUM	PROJ_NAME	EMP_NUM	EMP_NAME	JOB_CLASS	CHG_HOUR	HOURS
15	Evergreen	103	June E. Arbough	Elect. Engineer	84.50	23.8
15	Evergreen	101	John G. News	Database Designer	105.00	19.4
15	Evergreen	105	Alice K. Johnson *	Database Designer	105.00	35.7
15	Evergreen	106	William Smithfield	Programmer	35.75	12.6
15	Evergreen	102	David H. Senior	Systems Analyst	96.75	23.8
18	Amber Wave	114	Annelise Jones	Applications Designer	48.10	24.6
18	Amber Wave	118	James J. Frommer	General Support	18.36	45.3
18	Amber Wave	104	Anne K. Ramoras *	Systems Analyst	96.75	32.4
18	Amber Wave	112	Darlene M. Smithson	DSS Analyst	45.95	44.0
22	Rolling Tide	105	Alice K. Johnson	Database Designer	105.00	64.7
22	Rolling Tide	104	Anne K. Ramoras	Systems Analyst	96.75	48.4
22	Rolling Tide	113	Delbert K. Joenbrood *	Applications Designer	48.10	23.6
22	Rolling Tide	111	Geoff B. Wabash	Clerical Support	26.87	22.0
22	Rolling Tide	106	William Smithfield	Programmer	35.75	12.8
25	Starflight	107	Maria D. Alonzo	Programmer	35.75	24.6
25	Starflight	115	Travis B. Bawangi	Systems Analyst	96.75	45.8
25	Starflight	101	John G. News *	Database Designer	105.00	56.3
25	Starflight	114	Annelise Jones	Applications Designer	48.10	33.1
25	Starflight	108	Ralph B. Washington	Systems Analyst	96.75	23.6
25	Starflight	118	James J. Frommer	General Support	18.36	30.5
25	Starflight	112	Darlene M. Smithson	DSS Analyst	45.95	41.4

SOURCE: Course Technology/Cengage Learning

FIGURE 6.3 First normal form (1NF) dependency diagram



1NF (PROJ_NUM, EMP_NUM, PROJ_NAME, EMP_NAME, JOB_CLASS, CHG_HOURS, HOURS)

PARTIAL DEPENDENCIES:
(PROJ_NUM → PROJ_NAME)
(EMP_NUM → EMP_NAME, JOB_CLASS, CHG_HOUR)

TRANSITIVE DEPENDENCY:
(JOB_CLASS → CHG_HOUR)

SOURCE: Course Technology/Cengage Learning

2NF

- 2NF – *The Whole Key*
  - Remove partial dependencies
    - Break schema into two tables and use a bridge or associative table to create a relationship
  - **Project_Num** alone determines **Project_Name**
    - Don't need Emp_Num as part of the key
  - **Emp_Num** alone determines **Emp_Name**
    - Don't need Project_Num as part of the key
  - **Project(Project_Num, Project_Name)**

    Employee(Emp_Num, Emp_Name, Job_Class, Chg_Hour, Hours)

    Assignment(Project_Num, Emp_Num, Hours)

Second Normal Form (2NF)

PROJ_NUM	PROJ_NAME
15	Evergreen
18	Amber Wave
22	Rolling Tide
25	Starflight

PROJ_NUM	EMP_NUM	ASSIGN_HOURS
15	103	2.6
18	118	1.4
15	101	3.6
22	113	2.5
15	103	1.9
25	115	4.2
22	105	5.2
25	101	1.7
15	105	2
15	102	3.8
22	104	2.6
15	101	2.3
25	114	1.8
22	111	4
25	114	3.4
18	112	1.2
18	118	2
18	104	2.6
15	103	3
22	105	2.7
25	108	4.2
25	114	5.8
22	106	2.4

EMP_NUM	EMP_NAME		JOB_CLASS	CHG_HOUR
101	News	John	502	105
102	Senior	David	501	96.75
103	Arbough	June	503	84.5
104	Ramoras	Anne	501	96.75
105	Johnson	Alice	502	105
106	Smithfield	William	500	35.75
107	Alonzo	Maria	500	35.75
108	Washington	Ralph	501	96.75
109	Smith	Larry	501	96.75
110	Olenko	Gerald	505	22.56
111	Wabash	Geoff	506	26.87
112	Smithson	Darlene	507	45.95
113	Joenbrood	Delbert	508	48.1
114	Jones	Annelise	508	48.1
115	Bawangi	Travis	501	96.75
116	Pratt	Gerald	510	18.36
117	Williamson	Angie	509	15.56
118	Frommer	James	510	18.36

FIGURE 6.4 Second normal form (2NF) conversion results

Table name: PROJECT

PROJECT (PROJ_NUM, PROJ_NAME)

Table name: EMPLOYEE

EMPLOYEE (EMP_NUM, EMP_NAME, JOB_CLASS, CHG_HOUR)

TRANSITIVE DEPENDENCY
(JOB_CLASS → CHG_HOUR)

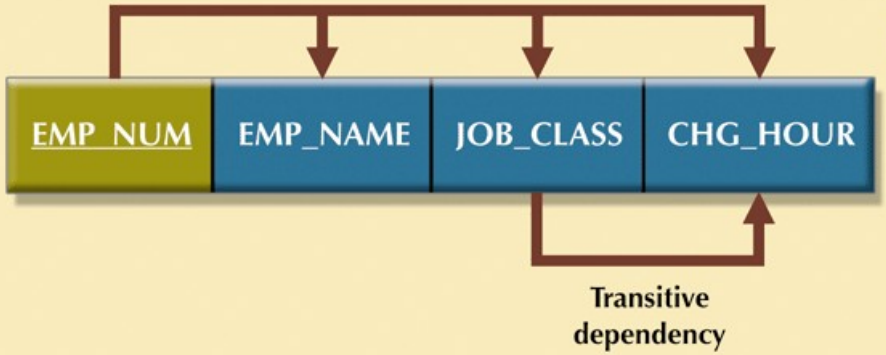


Table name: ASSIGNMENT

ASSIGNMENT (PROJ_NUM, EMP_NUM, ASSIGN_HOURS)

SOURCE: Course Technology/Cengage Learning

3NF

- 3NF – *And Nothing but the Key*
  - Remove transitive dependencies
    - Create a new table and establish a pk, fk relationship
  - Job_Class determines Chg_Hour
    - But Job_Class is not a primary key!
  - Create a fourth table so Chg_Hour is determined by the primary key
    - **Project(Project_Num, Project_Name)**
    - **Employee(Emp_Num, Emp_Name, Job_Class)**
    - **Assignment(Project_Num, Emp_Num, Hours)**
    - **Job(Job_Class, Chg_Hour)**

Third Normal Form (3NF)

PROJ_NUM	PROJ_NAME
15	Evergreen
18	Amber Wave
22	Rolling Tide
25	Starflight

PROJ_NUM	EMP_NUM	ASSIGN_HOURS
15	103	2.6
18	118	1.4
15	101	3.6
22	113	2.5
15	103	1.9
25	115	4.2
22	105	5.2
25	101	1.7
15	105	2
15	102	3.8
22	104	2.6
15	101	2.3
25	114	1.8
22	111	4
25	114	3.4
18	112	1.2
18	118	2
18	104	2.6
15	103	3
22	105	2.7
25	108	4.2
25	114	5.8
22	106	2.4

EMP_NUM	EMP_NAME		JOB_CLASS
101	News	John	502
102	Senior	David	501
103	Arbough	June	503
104	Ramoras	Anne	501
105	Johnson	Alice	502
106	Smithfield	William	500
107	Alonzo	Maria	500
108	Washington	Ralph	501
109	Smith	Larry	501
110	Olenko	Gerald	505
111	Wabash	Geoff	506
112	Smithson	Darlene	507
113	Joenbrood	Delbert	508
114	Jones	Annelise	508
115	Bawangi	Travis	501
116	Pratt	Gerald	510
117	Williamson	Angie	509
118	Frommer	James	510

Third Normal Form (3NF)

PROJ_NUM	EMP_NUM	ASSIGN_HOURS
15	101	3.6
25	101	1.7
15	101	2.3
15	102	3.8
15	103	2.6
15	103	1.9
15	103	3
22	104	2.6
18	104	2.6
22	105	5.2
15	105	2
22	105	2.7
22	106	2.4
25	108	4.2
22	111	4
18	112	1.2
22	113	2.5
25	114	1.8
25	114	3.4
25	114	5.8
25	115	4.2
18	118	1.4
18	118	2

JOB_CLASS	CHG_HOUR
500	35.75
501	96.75
502	105.00
503	84.50
504	67.90
505	55.78
506	26.87
507	45.95
508	48.10
509	34.55
510	18.36

FIGURE 6.5 Third normal form (3NF) conversion results

Table name: PROJECT

PROJECT (PROJ_NUM, PROJ_NAME)

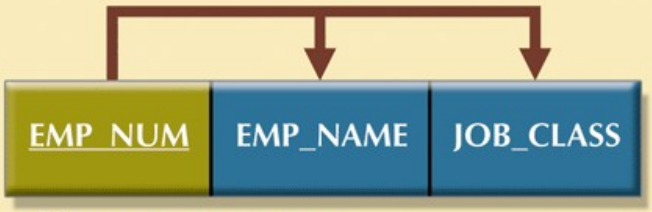


Table name: EMPLOYEE

EMPLOYEE (EMP_NUM, EMP_NAME, JOB_CLASS)

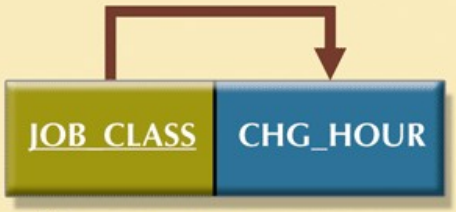


Table name: JOB

JOB (JOB_CLASS, CHG_HOUR)

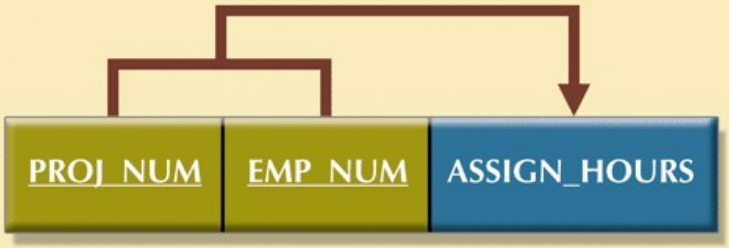


Table name: ASSIGNMENT

ASSIGNMENT (PROJ_NUM, EMP_NUM, ASSIGN_HOURS)

SOURCE: Course Technology/Cengage Learning