

Таблица до нормализации

ID	Employee	Supervisor	travel_required	Department	Project	Position
1	MICHAEL J. ANDERSON	NICOLE GRAY	yes	IT	Marketing optimization, demand forecasting	Senior Data Analyst
2	SCOTT G. ADAMS	CLARKE HAMBLEY	no	Tech	Automation of processes	Engineer
3	JAMES B. BARNES	MICHAEL J. ANDERSON	no	IT	Demand forecasting	Data Analyst
4	PAUL R. BROWN	MARGARET C. LEPAGE	no	Finance	Budget planning	Manager
5	JESSICA R. CURRIER	CATHERINE LEWIS	yes	HR	Hiring employees	Business partners
6	KRIS J. EIMICKE	MARGARET C. LEPAGE	yes	Finance	Budget planning	Account manager
7	STEVEN E. COPE	CATHERINE LEWIS	no	IT	Infrastructure development	Data engineer
8	REBECCA GREENFIELD	CLARKE HAMBLEY	yes	Tech	Automation of processes	Engineer
9	ANDREA L FRAVERT	MICHAEL J. ANDERSON	no	IT	Data visualization, Data analysis	Data Analyst
10	MATTHEW J. FOSTER	MICHAEL J. ANDERSON	no	IT	Sales optimization	Project manager

1 НФ

В базе данных не должно быть дубликатов и составных данных.

Почему таблица не в 1НФ? В атрибуте **Project** присутствуют списки.

Избавимся от них, приведя к простым.

ID	Employee	Supervisor	travel_required	Department	Project	Position
1	MICHAEL J. ANDERSON	NICOLE GRAY	yes	IT	Marketing optimization	Senior Data Analyst
2	MICHAEL J. ANDERSON	NICOLE GRAY	yes	IT	Demand forecasting	Senior Data Analyst
3	SCOTT G. ADAMS	CLARKE HAMBLEY	no	Tech	Automation of processes	Engineer
4	JAMES B. BARNES	MICHAEL J. ANDERSON	no	IT	Demand forecasting	Data Analyst
5	PAUL R. BROWN	MARGARET C. LEPAGE	no	Finance	Budget planning	Manager
6	JESSICA R. CURRIER	CATHERINE LEWIS	yes	HR	Hiring employees	Business partners
7	KRIS J. EIMICKE	MARGARET C. LEPAGE	yes	Finance	Budget planning	Account manager
8	STEVEN E. COPE	CATHERINE LEWIS	no	IT	Infrastructure development	Data engineer
9	REBECCA GREENFIELD	CLARKE HAMBLEY	yes	Tech	Automation of processes	Engineer
10	ANDREA L FRAVERT	MICHAEL J. ANDERSON	no	IT	Data visualization	Data Analyst
11	ANDREA L FRAVERT	MICHAEL J. ANDERSON	no	IT	Data analysis	Data Analyst
12	MATTHEW J. FOSTER	MICHAEL J. ANDERSON	no	IT	Sales optimization	Project manager

Также следует отметить, что **Employee** и **Supervisor** являются составными атрибутами и их нужно выносить в отдельные атрибуты **FirstName**, **SecondName**, **MiddleName**.

В данном случае,

во-первых, это не является критичным,

во-вторых, необходима актуализация данных, т.к **MiddleName** содержит сокращения,

а в-третьих, таблицу нужно предварительно нормализовать,

поэтому оставим эти атрибуты без изменений.

2 НФ

Необходимо устранить избыточные данные.

Почему таблица не в 2НФ? Обратим внимание, что значения **Employee** сопоставлено с несколькими значениями **Project**, а значения **Project** функционально не зависят от значений **Employee**. Следовательно, данное отношение не нормализовано до 2НФ.

ID	Employee	Supervisor	travel_required	Department	Position
1	MICHAEL J. ANDERSON	NICOLE GRAY	yes	IT	Senior Data Analyst
2	SCOTT G. ADAMS	CLARKE HAMBLEY	no	Tech	Engineer
3	JAMES B. BARNES	MICHAEL J. ANDERSON	no	IT	Data Analyst
4	PAUL R. BROWN	MARGARET C. LEPAGE	no	Finance	Manager
5	JESSICA R. CURRIER	CATHERINE LEWIS	yes	HR	Business partners
6	KRIS J. EIMICKE	MARGARET C. LEPAGE	yes	Finance	Account manager
7	STEVEN E. COPE	CATHERINE LEWIS	no	IT	Data engineer
8	REBECCA GREENFIELD	CLARKE HAMBLEY	yes	Tech	Engineer
9	ANDREA L FRAVERT	MICHAEL J. ANDERSON	no	IT	Data Analyst
10	MATTHEW J. FOSTER	MICHAEL J. ANDERSON	no	IT	Project manager

ID	Employee	Project
1	MICHAEL J. ANDERSON	Marketing optimization
2	MICHAEL J. ANDERSON	Demand forecasting
3	SCOTT G. ADAMS	Automation of processes
4	JAMES B. BARNES	Demand forecasting
5	PAUL R. BROWN	Budget planning
6	JESSICA R. CURRIER	Hiring employees
7	KRIS J. EIMICKE	Budget planning
8	STEVEN E. COPE	Infrastructure development
9	REBECCA GREENFIELD	Automation of processes
10	ANDREA L FRAVERT	Data visualization
11	ANDREA L FRAVERT	Data analysis
12	MATTHEW J. FOSTER	Sales optimization

3 НФ

Необходимо устранить данные, не зависящие от ключа.

Почему таблица не в 3НФ? Несложно заметить, что от **Employee** транзитивно зависит **Supervisor**, следовательно выносим этот атрибут в отдельную таблицу.

ID	Employee	travel_required	Department	Position
1	MICHAEL J. ANDERSON	yes	IT	Senior Data Analyst
2	SCOTT G. ADAMS	no	Tech	Engineer
3	JAMES B. BARNS	no	IT	Data Analyst
4	PAUL R. BROWN	no	Finance	Manager
5	JESSICA R. CURRIER	yes	HR	Business partners
6	KRIS J. EIMICKE	yes	Finance	Account manager
7	STEVEN E. COPE	no	IT	Data engineer
8	REBECCA GREENFIELD	yes	Tech	Engineer
9	ANDREA L FRAVERT	no	IT	Data Analyst
10	MATTHEW J. FOSTER	no	IT	Project manager

ID	Employee	Supervisor
1	MICHAEL J. ANDERSON	NICOLE GRAY
2	SCOTT G. ADAMS	CLARKE HAMBLEY
3	JAMES B. BARNS	MICHAEL J. ANDERSON
4	PAUL R. BROWN	MARGARET C. LEPAGE
5	JESSICA R. CURRIER	CATHERINE LEWIS
6	KRIS J. EIMICKE	MARGARET C. LEPAGE
7	STEVEN E. COPE	CATHERINE LEWIS
8	REBECCA GREENFIELD	CLARKE HAMBLEY
9	ANDREA L FRAVERT	MICHAEL J. ANDERSON
10	MATTHEW J. FOSTER	MICHAEL J. ANDERSON

Помимо этого, **Department** имеет прямую зависимость от **Employee**, но косвенную от **Position**, следовательно выносим его в отдельную таблицу.

ID	Position	travel_required	Employee
1	Senior Data Analyst	yes	MICHAEL J. ANDERSON
2	Engineer	no	SCOTT G. ADAMS
3	Data Analyst	no	JAMES B. BARNS
4	Manager	no	PAUL R. BROWN
5	Business partners	yes	JESSICA R. CURRIER
6	Account manager	yes	KRIS J. EIMICKE
7	Data engineer	no	STEVEN E. COPE
8	Engineer	yes	REBECCA GREENFIELD
9	Data Analyst	no	ANDREA L FRAVERT
10	Project manager	no	MATTHEW J. FOSTER

Необходимо отметить, что в данной таблице присутствует потенциальный первичный составной ключ **Position - travel_required**.
Почему именно он?
Потому что,
- во-первых, **travel_required** - это признак, будут ли командировки у данной позиции/вакансии/ставки,
- а во-вторых, **travel_required** напрямую зависит от **Position** и это видно по **Engineer**.

ID	Employee	Supervisor
1	MICHAEL J. ANDERSON	NICOLE GRAY
2	SCOTT G. ADAMS	CLARKE HAMBLEY
3	JAMES B. BARNS	MICHAEL J. ANDERSON
4	PAUL R. BROWN	MARGARET C. LEPAGE
5	JESSICA R. CURRIER	CATHERINE LEWIS
6	KRIS J. EIMICKE	MARGARET C. LEPAGE
7	STEVEN E. COPE	CATHERINE LEWIS
8	REBECCA GREENFIELD	CLARKE HAMBLEY
9	ANDREA L FRAVERT	MICHAEL J. ANDERSON
10	MATTHEW J. FOSTER	MICHAEL J. ANDERSON

ID	Employee	Department
1	MICHAEL J. ANDERSON	IT
2	SCOTT G. ADAMS	Tech
3	JAMES B. BARNS	IT
4	PAUL R. BROWN	Finance
5	JESSICA R. CURRIER	HR
6	KRIS J. EIMICKE	Finance
7	STEVEN E. COPE	IT
8	REBECCA GREENFIELD	Tech
9	ANDREA L FRAVERT	IT
10	MATTHEW J. FOSTER	IT

ИТОГО: 1НФ + 2НФ + 3НФ

ID	Employee	Project
1	MICHAEL J. ANDERSON	Marketing optimization
2	MICHAEL J. ANDERSON	Demand forecasting
3	SCOTT G. ADAMS	Automation of processes
4	JAMES B. BARNS	Demand forecasting
5	PAUL R. BROWN	Budget planning
6	JESSICA R. CURRIER	Hiring employees
7	KRIS J. EIMICKE	Budget planning
8	STEVEN E. COPE	Infrastructure development
9	REBECCA GREENFIELD	Automation of processes
10	ANDREA L FRAVERT	Data visualization

11	ANDREA L FRAVERT	Data analysis
12	MATTHEW J. FOSTER	Sales optimization

ID	Position	travel_required	Employee
1	Senior Data Analyst	yes	MICHAEL J. ANDERSON
2	Engineer	no	SCOTT G. ADAMS
3	Data Analyst	no	JAMES B. BARNS
4	Manager	no	PAUL R. BROWN
5	Business partners	yes	JESSICA R. CURRIER
6	Account manager	yes	KRIS J. EIMICKE
7	Data engineer	no	STEVEN E. COPE
8	Engineer	yes	REBECCA GREENFIELD
9	Data Analyst	no	ANDREA L FRAVERT
10	Project manager	no	MATTHEW J. FOSTER

ID	Employee	Supervisor
1	MICHAEL J. ANDERSON	NICOLE GRAY
2	SCOTT G. ADAMS	CLARKE HAMBLEY
3	JAMES B. BARNS	MICHAEL J. ANDERSON
4	PAUL R. BROWN	MARGARET C. LEPAGE
5	JESSICA R. CURRIER	CATHERINE LEWIS
6	KRIS J. EIMICKE	MARGARET C. LEPAGE
7	STEVEN E. COPE	CATHERINE LEWIS
8	REBECCA GREENFIELD	CLARKE HAMBLEY
9	ANDREA L FRAVERT	MICHAEL J. ANDERSON
10	MATTHEW J. FOSTER	MICHAEL J. ANDERSON

ID	Employee	Department
1	MICHAEL J. ANDERSON	IT
2	SCOTT G. ADAMS	Tech
3	JAMES B. BARNS	IT
4	PAUL R. BROWN	Finance
5	JESSICA R. CURRIER	HR
6	KRIS J. EIMICKE	Finance
7	STEVEN E. COPE	IT
8	REBECCA GREENFIELD	Tech
9	ANDREA L FRAVERT	IT
10	MATTHEW J. FOSTER	IT

Дополнительно

Нельзя не отметить, что значения в атрибутах **Employee** и **Supervisor** имеют похожие значения. Выглядит логичным создать отдельную таблицу **User/Person**, в которой будут все Сотрудники и Руководители, т.к. независимо ни от чего, для нас они все - сотрудники компании.

В таком случае тип сотрудника мы вполне могли бы определять как отдельный атрибут.

Но в данном учебном примере от нас этого не требуется, поэтому оставим все как есть.