Дана таблица, содержащая данные о сотрудниках, такие как:

* **Employee** – Ф.И.О сотрудника;
* **Supervisor** – Ф.И.О. начальника (руководителя);
* **Travel\_require**d – связана ли работа с разъездами;
* **Department** – отдел, департамент в котором работает сотрудник;
* **Project** – проекты, в которых принимает участие сотрудник;
* **Position** – занимаемая должность.

| **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. BARNS | 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 |

Необходимо произвести нормализацию таблицы, объясняя на каждом этапе к какой нормальной форме производится преобразование.

Решение:

1. **На первом этапе приведем таблицу к 1НФ**. Таблица находится в 1НФ тогда, когда каждая строка содержит только одно значение для каждого из атрибутов. В строках 1 и 9 атрибут «Project» содержит по два наименования проектов, в которых принимают участие сотрудники. Для приведения к 1НФ необходимо каждую из строк 1 и 9 разбить на две, в результате получится следующая таблица:

Таблица 1

| **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. BARNS | 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 |

1. **На втором этапе приведем таблицу ко 2НФ**. Таблица находится во 2НФ тогда, когда она находится в 1НФ и каждый неключевой атрибут (неприводимо) функционально полно зависит от ее потенциального ключа, т.е. если ключ является составным, то атрибут зависит от всего ключа и не зависит от его частей. Единственный потенциальный ключ содержит атрибуты «Employee» и «Project». Т.к. все неключевые атрибуты зависят только от части потенциального ключа (атрибута «Employee»), то таблица не находится во 2НФ и для приведения ко 2НФ требуется ее преобразование. В результате получим следующие две таблица, каждая из которых находится во 2НФ (в первой таблице ключевым атрибутом является «ID», во второй – атрибут «Employee»:

Таблица 2.1

| **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 |

Таблица 2.2

| **Employee** | **Supervisor** | **Travel\_ required** | **Department** | **Position** |
| --- | --- | --- | --- | --- |
| MICHAEL J. ANDERSON | NICOLE GRAY | yes | IT | Senior Data Analyst |
| SCOTT G. ADAMS | CLARKE HAMBLEY | no | Tech | Engineer |
| JAMES B. BARNS | MICHAEL J. ANDERSON | no | IT | Data Analyst |
| PAUL R. BROWN | MARGARET C. LEPAGE | no | Finance | Manager |
| JESSICA R. CURRIER | CATHERINE LEWIS | yes | HR | Business partners |
| KRIS J. EIMICKE | MARGARET C. LEPAGE | yes | Finance | Account manager |
| STEVEN E. COPE | CATHERINE LEWIS | no | IT | Data engineer |
| REBECCA GREENFIELD | CLARKE HAMBLEY | yes | Tech | Engineer |
| ANDREA L FRAVERT | MICHAEL J. ANDERSON | no | IT | Data Analyst |
| MATTHEW J. FOSTER | MICHAEL J. ANDERSON | no | IT | Project manager |

1. **На третьем этапе приведем таблицы к 3НФ**. Для того, чтобы таблица находилась в 3 НФ необходимо, чтобы она находилась во 2НФ и отсутствовали транзитивные зависимости неключевых атрибутов от ключевых. В Таблице 2.2 атрибуты «Department» и «Supervisor» транзитивно зависят от ключевого атрибута «Employee» (через неключевой атрибут «Position»). Для приведения к 3НФ преобразуем таблицу 2.2 так, чтобы в получившихся таблицах все неключевые атрибуты зависели только от ключевого. В результате получим следующий набор таблиц:

Таблица 3.1

| **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 |

Таблица 3.2

| **Employee** | **Travel\_ required** | **Position** |
| --- | --- | --- |
| MICHAEL J. ANDERSON | yes | Senior Data Analyst |
| SCOTT G. ADAMS | no | Engineer |
| JAMES B. BARNS | no | Data Analyst |
| PAUL R. BROWN | no | Manager |
| JESSICA R. CURRIER | yes | Business partners |
| KRIS J. EIMICKE | yes | Account manager |
| STEVEN E. COPE | no | Data engineer |
| REBECCA GREENFIELD | yes | Engineer |
| ANDREA L FRAVERT | no | Data Analyst |
| MATTHEW J. FOSTER | no | Project manager |

Таблица 3.3

| **Position** | **Department** | **Supervisor** |
| --- | --- | --- |
| Senior Data Analyst | IT | NICOLE GRAY |
| Engineer | Tech | CLARKE HAMBLEY |
| Data Analyst | IT | MICHAEL J. ANDERSON |
| Manager | Finance | MARGARET C. LEPAGE |
| Business partners | HR | CATHERINE LEWIS |
| Account manager | Finance | MARGARET C. LEPAGE |
| Data engineer | IT | CATHERINE LEWIS |
| Project manager | IT | MICHAEL J. ANDERSON |

В таблице 3.1 ключевым атрибутом является «**ID**», в таблице 3.2 – «**Employee**», в таблице 3.3. – «**Position**».

Ниже приводится схема базы данных, составленной из полученных таблиц:

