**Normalização da Base de Dados**

O processo de normalização, é um passo fundamental para a construção de um sistema de base de dados rigoroso.

Para isso, resolvemos fazer um povoamento *ad hoc* (feito numa folha de cálculo) das tabelas da nossa base de dados para nos ajudar a seguir os sete passos do processo de normalização (*Connolly, T., Begg, C., Database Systems, A Practical Approach to Design, Implementation, and Management, Addison-Wesley, 4ª Edição, 2004).*

1. **Unormalized Form (UNF)**

Ao chegarmos ao processo de normalização, tínhamos já concluído o modelo conceptual do nosso sistema, bem como um primeiro esboço do modelo lógico que foi alvo de um povoamento *ad hoc*, cujo resultado se apresenta de seguida (colunas a vermelho correspondem a PK e a azul a FK):

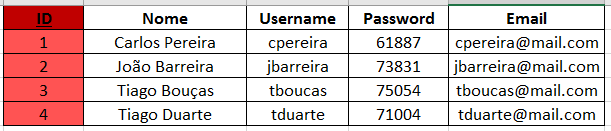
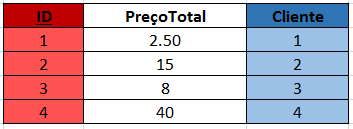


Tabela 1 - Cliente

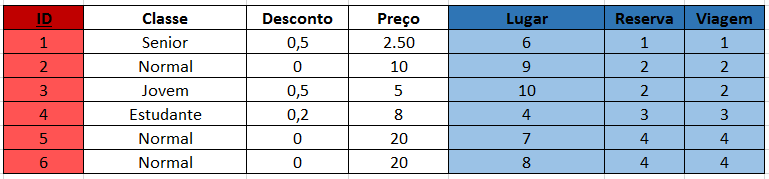
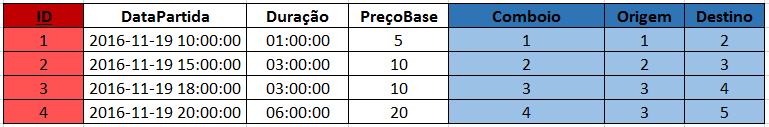


Tabela 4 - Viagem

Tabela 3 - Bilhete

Tabela 2 - Reserva

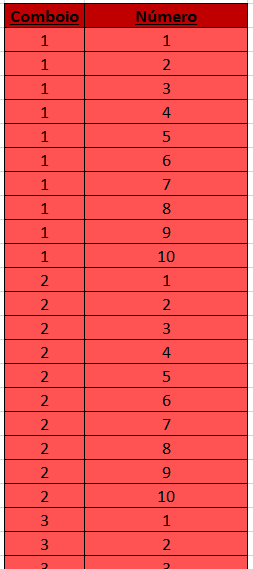


Tabela 5 - Estação

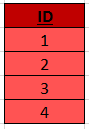


Tabela 6 - Comboio

Tabela 7 - Lugar (não está na imagem: 10 lugares tanto para o comboio 3 como para o 4)

1. **First Normal Form (1NF)**

* **A interseção entre cada linha e cada coluna (i.e. cada célula) deve conter um e um só valor.**

O nosso esquema inicial já obedecia a esta regra, pelo que não foi necessário alterar nada.

1. **Second Normal Form (2NF)**

* **Todos os atributos não primários de uma tabela devem ser funcionalmente dependentes da chave primária dessa tabela**.

Todas as tabelas do nosso esquema inicial obedecem a esta regra:

* + **Cliente** – *nome*, *username*, *password* e *email* são atributos funcionalmente dependentes de cada cliente (ID).
  + **Reserva** – *preço total* e *cliente* são atributos funcionalmente dependentes de cada reserva (ID).
  + **Bilhete** – *classe*, *desconto*, *preço*, *lugar*, *reserva* e *viagem* são atributos funcionalmente dependentes de cada bilhete (ID).
  + **Viagem** – *data de partida*, *duração*, *preço* *base*, *comboio*, *origem* e *destino* são atributos funcionalmente dependentes de cada viagem (ID).
  + **Estação** – *nome* e *país* são atributos funcionalmente dependentes de cada estação (ID).
  + **Comboio** –esta tabela não tem quaisquer atributos que não a PK.
  + **Lugar** –esta tabela não tem quaisquer atributos não primários.

1. **Third Normal Form (3NF)**

* **Nenhum atributo não primário deve ser transitivamente dependente em relação à chave primária (i.e. todas as colunas de uma tabela devem apenas poder ser determinadas pela coluna da PK e não por qualquer outra coluna da tabela).**

A única tabela que desobedecia a esta critério era a tabela *Bilhete*, visto que a coluna relativa ao *desconto* associado à *classe* do bilhete, deduzia-se pela classe do bilhete e não pelo bilhete em si. Isto é, a coluna *desconto* era transitivamente dependente em relação à chave primária.

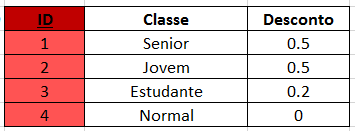
O que resolvemos fazer foi eliminar a coluna *desconto*, criando uma nova tabela para o efeito chamada *Desconto* em que apenas se associava o nome da classe (e.g. Sénior, Estudante, etc) ao seu desconto. Mudanças essas que podemos ver nas duas tabelas seguintes:

Tabela 3 (v2) - Bilhete

Tabela 8 - Desconto

1. **Boyce-Codd Normal Form (BCNF)**

* **Todos os determinantes (i.e. atributo, simples ou composto, em relação ao qual algum outro atributo está funcionalmente dependente) de uma tabela devem ser chaves candidatas dessa mesma tabela.**

Apresentamos, de seguida, a lista das chaves candidatas de cada tabela e as:

* + **Cliente**:
    - Chaves candidatas: {ID}, {Username}, {Email}
    - Relações de dependência:
      * {ID} -> {nome, username, password, email}
      * {Username} -> {ID, nome, password, email}
      * {Email} -> {ID, nome, username, password}
  + **Reserva:**
    - Chave candidata: {ID}
    - Relações de dependência:
      * {ID} -> {preço total, cliente}
  + **Bilhete:**
    - Chaves candidatas: {ID}, {Lugar, Viagem}
    - Relações de dependência:
      * {ID} -> {classe, preço, lugar, reserva, viagem}
      * {Lugar, Viagem} -> {
      * {Classe, Viagem} -> {preço}
  + **Desconto:**
    - {ID}
    - {Classe}
  + **Viagem:**
    - {ID}
    - {Data partida, Comboio}
  + **Estação:**
    - {ID}
    - {Nome}
  + **Comboio:**
    - {ID}
  + **Lugar:**
    - {Comboio, Número}

Como todas as relações da dependência das tabelas têm por base atributos que são chaves candidatas, o nosso modelo não desobedece aos requerimentos do BCNF.

1. **Fourth Normal Form (4NF)**
2. **Fifth Normal Form (5NF)**