Sheet: airports

|  |  |  |  |
| --- | --- | --- | --- |
| **Cột** | **Kiểu** | **Mô tả** | **Phương pháp fill dữ liệu** |
| AirportID | int | ID duy nhất cho sân bay | Dùng trực tiếp từ file (PK) |
| Name | str | Tên sân bay (VD: Tan Son Nhat Intl) | Giữ nguyên, strip khoảng trắng |
| City | str | Thành phố | Capitalize chữ cái đầu |
| Country | str | Quốc gia | Chuẩn hóa tên quốc gia (VD: “Viet Nam” → “Vietnam”) |
| IATA (International Air  Transport Association) | str | 3-letter airport code | Nếu rỗng → fill = “N/A” |
| ICAO (International Civil  Aviation Organization) | str | 4-letter code | Tương tự |
| Latitude, Longitude | float | Toạ độ địa lý | Chuyển về float, drop nếu null |
| Altitude | float | Độ cao (feet) | Chuyển thành mét nếu cần |
| Timezone | float | UTC offset | Giữ nguyên |
| DST, TZ Database Time Zone, Type, Source | str | Metadata | Dùng nguyên bản hoặc drop nếu phân tích không cần |

Primary Key: airport\_id

Sheet: airlines

|  |  |  |  |
| --- | --- | --- | --- |
| **Cột** | **Kiểu** | **Mô tả** | **Phương pháp fill dữ liệu** |
| AirlineID | int | ID duy nhất cho hãng bay | Giữ nguyên (PK) |
| Name | str | Tên hãng bay (Vietnam Airlines) | Clean text |
| Alias | str | Tên khác (ví dụ viết tắt) | Fill “N/A” nếu rỗng |
| IATA, ICAO | str | Code hãng | Fill “N/A” nếu rỗng |
| Callsign | str | Dấu hiệu liên lạc radio | Giữ nguyên hoặc drop |
| Country | str | Quốc gia đăng ký | Chuẩn hóa tên quốc gia |
| Active | str | ‘Y’ hoặc ‘N’ | Convert sang boolean True/False |

Primary Key: airline\_id

Sheet: routes

|  |  |  |  |
| --- | --- | --- | --- |
| **Cột** | **Kiểu** | **Mô tả** | **Phương pháp fill dữ liệu** |
| AirlineID | int | Liên kết tới airlines | Check tồn tại trong airlines, drop nếu không có |
| DepartureAirportID | int | ID sân bay đi | Check tồn tại trong airports |
| DestinationAirportID | int | ID sân bay đến | Check tồn tại trong airports |
| Stops | int | Số điểm dừng | Fill 0 nếu null |
| Codeshare, Equipment | str | Metadata | Có thể giữ nguyên hoặc drop |