**Giải thích dữ liệu**

Tên của bộ dữ liệu : Weather Conditions in World War Two.

Link của bộ dữ liệu: <https://www.kaggle.com/smid80/weatherww2/data>

Giải thích loại bỏ dữ liệu thừa:

|  |  |  |
| --- | --- | --- |
| Tên thuộc tính ban đầu | Ý nghĩa | Lí do loại bỏ |
| WindGustSpd | Tốc độ gió cực đại với đơn vị đo là Km/h | Trùng lặp ý nghĩa và giá trị với thuộc tính SPD, ở đây SPQ là tốc độ gió cực đại với đơn vị đo là Hải lí |
| Snowfall | Độ dày của tuyết rơi với đơn vị đo là mm | Trùng lặp ý nghĩa và giá trị với thuộc tính SNF, ở đây SNF là độ dày của tuyết rơi với đơn vị đo là inch |
| MeanTemp | Nhiệt độ trung bình theo độ C | Do đã có maxtemp (nhiệt độ lớn nhất) và mintemp (nhiệt độ nhỏ nhất) nên không cần thiết thuộc tính meantemp |
| YR | Năm quan sát | Trùng lặp với thuộc tính Data(ngày tháng năm) quan sát |
| MO | Tháng quan sát |
| DA | Ngày quan sát |
| MAX | Nhiệt độ lớn nhất theo độ F | Trùng lặp ý nghĩa với Maxtemp và mintemp, ở đây maxtemp và mintemp là đo theo độ C |
| MIN | Nhiệt độ nhỏ nhất theo độ F |
| MEA | Nhiệt độ trung bình theo độ F |
| FT | Frozen Ground Top (depth in inches) | Không có dữ liệu. |
| FB | Frozen Ground Base (depth in inches) |
| FTI | Frozen Ground Thickness (thickness in inches) |
| ITH | Ice Thickness on Water (inches and tenths) |
| SD3 | Snow depth at 0030 GMT includes ice pellets in inches and tenths |
| RHX | 24-hour maximum relative humidity, as a whole percent |
| RHN | 24-hour minimum relative humidity, as a whole percent |
| RVG | River guage in feet and tenths |
| WTE | Water equivalent of snow and ice on ground in inches and hundredths |
| TSHDSBRSGF | Là viết tắt của: Thunder; Sleet; Hail; Dust or Sand; Smoke or Haze; Blowing Snow; Rain; Snow; Glaze; Fog; 0 = No, 1 = Yes | Trùng lặp ý nghĩa và giá trị với thuộc tính PoorWeather |
| SND | Snow depth (includes ice pellets) recorded at 1200 GMT except 0000 GMT in Far East Asian Area in inches and tenths | Do chỉ có một giá trị duy nhất là 0, không có ý nghĩa để sử dụng |
| PRCP | Lượng mưa theo đơn vị là inch và lấy 2 chữ số thập phân | Trùng ý nghĩa và giá trị với thuộc tính Precip, ở đây nghĩa là lượng mưa theo mm |

Giải thích thuộc tính bảng fact:

|  |  |  |
| --- | --- | --- |
| Tên thuộc tính | Ý nghĩa | Ghi chú |
| FactID | Khóa chính |  |
| StationID | Khóa ngoại |  |
| DateID | Khóa ngoại |  |
| PoorWeatherID | Khóa ngoại |  |
| Precip | Lượng mưa theo đơn vị mm |  |
| MaxCelsius | Nhiệt độ tối đa (độ C) | Đổi tên từ thuộc tính MaxTemp |
| MinCelsius | Nhiệt độ tối thiểu (độ C) | Đổi tên từ thuộc tính MinTemp |
| MaximumWindySpeedKnots | Tốc độ gió cực đại theo hải lí | Đổi tên từ thuộc tính SPD |
| SnowDepthInch | Độ dày của tuyết theo đơn vị inch | Đổi tên từ thuộc tính SNF |
| WindyTime | Thời điểm gió thổi | Đổi tên từ thuộc tính PGT |
| WindyDriection | Hướng gió thổi | Đổi tên từ thuộc tính DR |

Giải thích thuộc tính bảng Date:

|  |  |
| --- | --- |
| Tên thuộc tính | Ý nghĩa |
| DateID | Khóa chính |
| Day | Ngày quan sát |
| Mon | Tháng quan sát |
| Year | Năm quan sát |

Giải thích thuộc tính bảng Country

|  |  |
| --- | --- |
| Tên thuộc tính | Ý nghĩa |
| CountryID | Khóa chính |
| CountryName | Tên của quốc gia |

Giải thích thuộc tính bảng WeatherStation

|  |  |
| --- | --- |
| Tên thuộc tính | Ý nghĩa |
| StationID | Khóa chính |
| StationName | Tên của trạm đo thời tiết |

Giải thích bảng PoorWeather

|  |  |  |
| --- | --- | --- |
| Tên thuộc tính | Ý nghĩa | Ghi chú |
| PoorWeatherID | Khóa chính |  |
| Thunder | Thời tiết có sét đánh | Có 3 giá trị thể hiện  1 : có thời tiết tương ứng xảy ra  0 : không có thời tiết ứng xảy ra  -1 : không xác định |
| Sleet | Thời tiết có mưa đá |
| Hail | Thời tiết có gió thổi mạnh |
| DustOrSand | Thời tiết có bụi hoặc cát |
| SmokeOrHaze | Thời tiết có khói hoặc sương mù |
| BlowingSnow | Thời tiết có tuyết kèm theo gió thổi |
| Rain | Thời tiết có mưa |
| Snow | Thời tiết có tuyết rơi |
| Glaze |  |
| Fog | Thời tiết có sương mù vào buổi sáng |