إشكالية

تعد مشكلة تحديد ازدحام المرور التي نعاني منها في عصرنا و نبحث دائما عن حلول لتفادي الطرق المزدحمة

مقدمة

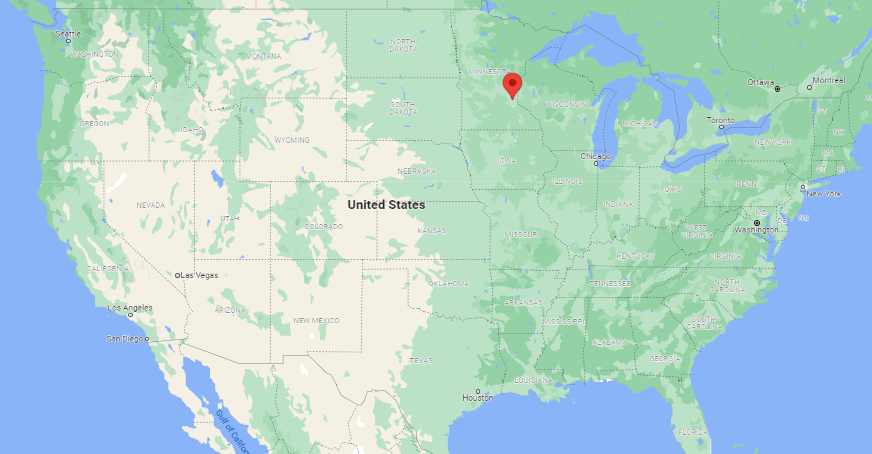
من اجل هذا سنقوم بانشاء برنامج باستخدام تعلم الالة حتا نحدد كمية حركة المرور من مجموعة من الامؤثرات التي تحصلنا عليها من قاعدة بيانات لطريق بين

Minneapolis

and

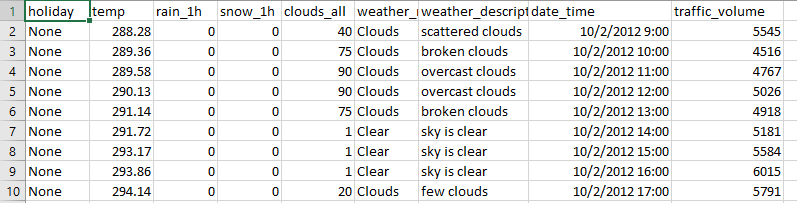
Saint Paul

In USA



هتان مدينتان متجاورتان في الولايات المتحدة

تحصلنا عل قاعدة بيانات لاكنها تحتاج ال معالجة



++++++++++++++++

Holyday

|  |  |
| --- | --- |
| value | duplicated |
| None | 48143 |
| Columbus Day | 5 |
| Veterans Day | 5 |
| Thanksgiving Day | 6 |
| Christmas Day | 6 |
| New Years Day | 6 |
| Washingtons Birthday | 5 |
| Memorial Day | 5 |
| Independence Day | 5 |
| State Fair | 5 |
| Labor Day | 7 |
| Martin Luther King Jr Day | 6 |
|  |  |

تمثل الأعياد الوطنية في الولايات المتحدة

من الناحية الإحصائية نجد ان من الخطء استخدام هذا العمود لانه يحتوي عل العديد من الأخطاء في التواريخ

بالإضافة ال ان نسبة القيم الى none

تساوي 0.001

و هيا تعتبر قيم شاذة

++++++++++++++++

Temp

تمثل درجة الحرارة تما قياسها بوحدة الكالفن

تعريف الكالفن

سنقون بتحويلها ال السيليسيوس من اجل دراستها باستخدام هذا الكود



لاحضال وجود قيم outline

درجة حرارة 270 درجة تحت الصفر و هذا بالتأكيد خطء في القياس

و لحذفها



rain\_1h ماينقصها والو

++++++++++++++++

snow\_1h ماينقصها والو

++++++++++++++++

clouds\_all ماينقصها والو

++++++++++++++++

weather

قيم نصية و يمكن استنتاجها ببساطة من rain\_1h snow\_1h clouds\_all

قمنا بحذف هذا العمود

++++++++++++++++

weather\_main

قيم نصية و يمكن استنتاجها ببساطة من rain\_1h snow\_1h clouds\_all

قمنا بحذف هذا العمود

++++++++++++++++

date\_time

يحتوي عل قيم نصية لتاريخ و الوقت

حتا نستطيع التعامل معه كان علينا تقسيمه ال أجزاء

قمنا بتحويل النص ال قيم للوقت

و استخرجنا

السنة

الشهر / 0 جانفي

1 فيفري

اليوم من الشهر

اليوم من الأسبوع : الاثنين يمثل 0 الثلثاء 1 .............................................. codification

الساعة

الدقائق كلها اصفار لا حاجة لاستخراجها

---++++++

بعد تنفيذ الكود تحصلنا عل

