|  |
| --- |
| Seattle Call Data |

## Introducción

Este dataset representa la actividad de respuesta policial. Cada fila es un registro de un Call for Service (CfS) registrado en el Centro de Comunicaciones del Departamento de Policía de Seattle (SPD). Las llamadas se originaron en la comunidad y van desde en curso o emergencias activas hasta solicitudes de resolución de problemas. Además, los oficiales registrarán las llamadas de sus observaciones del campo.

Actualmente el sistema asigna prioridades a las llamadas de una forma arbitraria, lo cual esta ocasionando un uso ineficiente de los recursos. Creando confusión en la población y la falta de credibilidad del SDP para accionar ante posibles crímenes.

## Objetivo

El SDP se encuentra en una transformación digital y sus proyectos están dirigiéndose hacia Big Data y ML. Desean crear una arquitectura de Big Data que les permita realizar análisis descriptivo sobre diferentes fuentes de información, buscar correlaciones y poder contar con modelos de ML que puedan tener una predicción sobre el tipo de llamada para poder asignar una prioridad real y así ser mas eficientes y recuperar su credibilidad.

## Datasets

* Call Data – Llamadas realizadas hacia el SDP
  + <https://data.seattle.gov/Public-Safety/Call-Data/33kz-ixgy>
* Crime Data – Contiene los crímenes denunciados al SDP
  + <https://data.seattle.gov/Public-Safety/Crime-Data/4fs7-3vj5>
* Road Weather Information Stations – Cada estación tiene un sensor de temperatura que mide la temperatura de la superficie de la calle y un sensor que mide la temperatura del aire ambiente en la estación cada segundo.
  + <https://data.seattle.gov/Transportation/Road-Weather-Information-Stations/egc4-d24i>
* City Sponsored Events Simple – Contiene la información de los eventos que se realizan en la ciudad.
  + <https://data.seattle.gov/Community/City-Sponsored-Events-Simple/39xq-zv8h>

## Hipótesis

El SDP realizo una lluvia de ideas entre sus diferentes directores para plantear las siguientes hipótesis

* ¿El clima es un factor importante para determinar si se cometerán crímenes?
  + ¿En qué zonas?
  + ¿Tipos de Crímenes?
  + ¿Existen más llamadas en función de una estación del año determinada?
* ¿Los eventos en la ciudad generan alta probabilidad de crímenes?
  + ¿Qué tipo de eventos?
  + ¿En qué zonas?
  + ¿Qué temporada del año?

## Productos finales

* Diseñar e implementar una arquitectura de Big Data que permita almacenar y procesar diferentes Datasets.
* Realizar un Data Discovery entre los Datasets y presentarlos en una Infografía.
* Dashboard interactivo que permita describir la situación actual correlacionando los Datasets.\*
* Desarrollar un modelo ML que permita a través de un API asignar la prioridad de la llamada.\*\*

\* Crédito Extra realizar un Dashboard tipo producto (publicarlo en un Web Site)

\*\*Crédito extra realizar la publicación del API