

Datathon reto

Por: Sebastián Muro Sánchez

Email: sebastian.muro@train.ia.center

Utilizando el [Diccionario de Datos de Estadística de Producción Agrícola 2003-Presente](#) de la SIAP es posible identificar las columnas relevantes para la tarea.

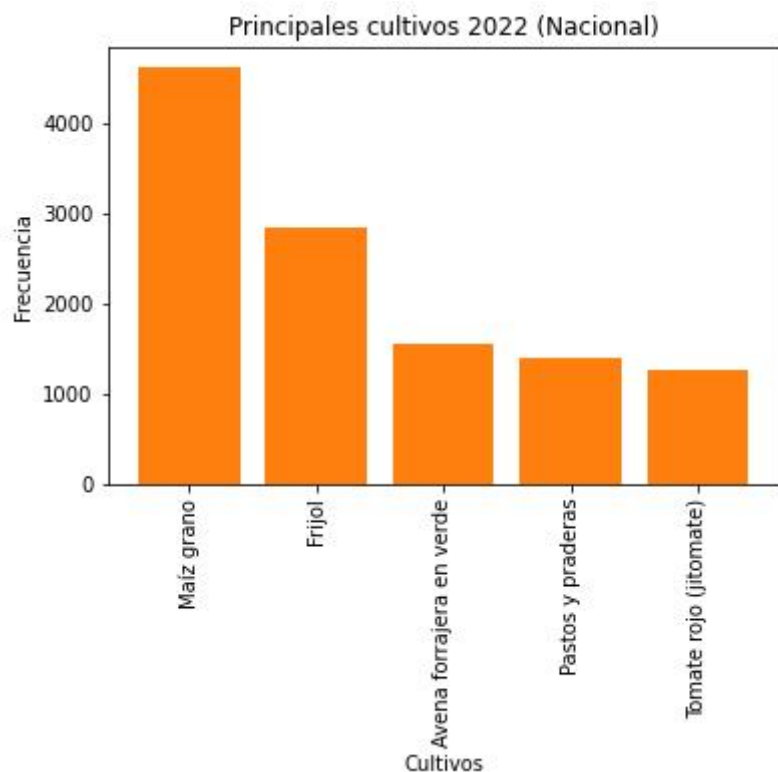
Para los análisis estatales y municipales es necesario tomar referencia de las columnas "Idestado" que muestra el número único de referencia de cada estado, "Nomestado" que muestra el nombre del estado; y de "Idmunicipio" y "Nommunicipio" para los municipios. Para la identificación de los principales cultivos (en cuanto a frecuencia) se utilizan las columnas de "Idcultivo" y "Nomcultivo".

El notebook "merger_nb.ipynb" se encarga de unir todos los datasets individuales de cada año en una sola base de datos.

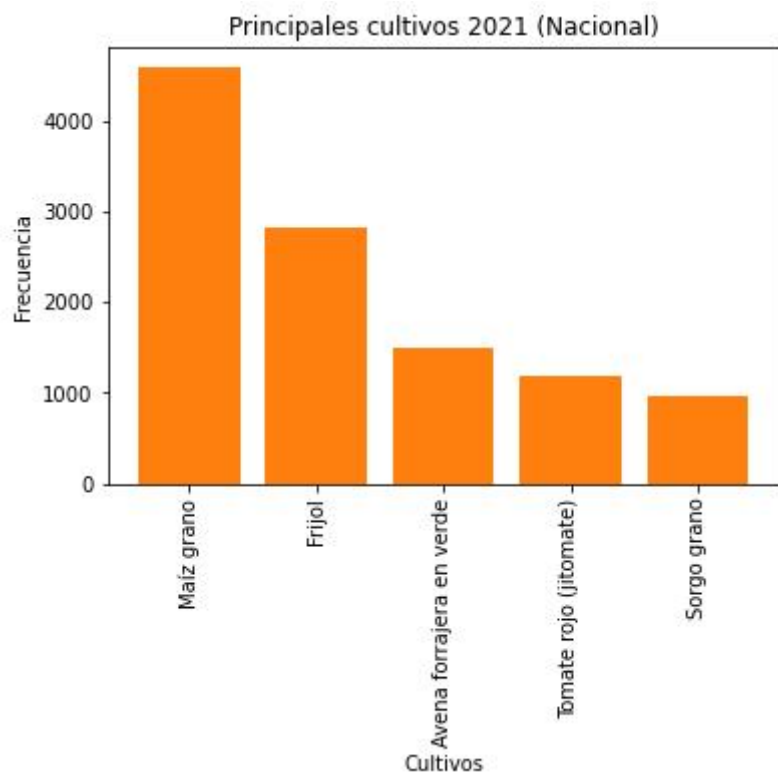
El notebook "agr_v2_nb.ipynb" analiza los datos nacionales, estatales o municipales. (Por razones de tiempo y limitantes de hardware no se realizó el análisis para todos los años, solo de 2020 a 2022).

Nacional:

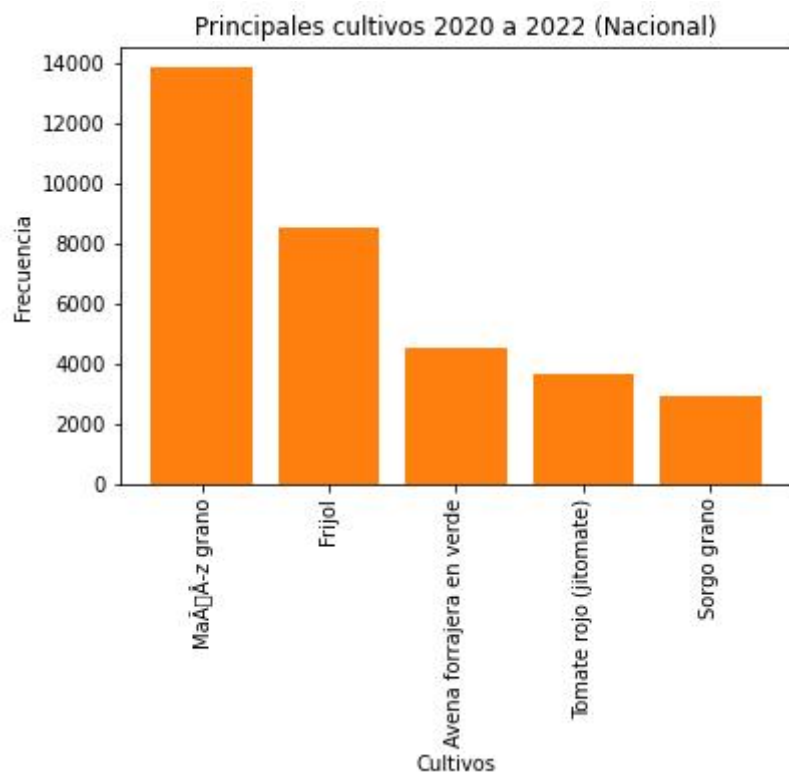
En el 2022 se produjeron 324 tipos de cultivos diferentes, de los cuales los principales fueron los siguientes:



En el 2021 se produjeron tan solo 82 tipos de cultivos diferentes, de los cuales los principales fueron los siguientes:



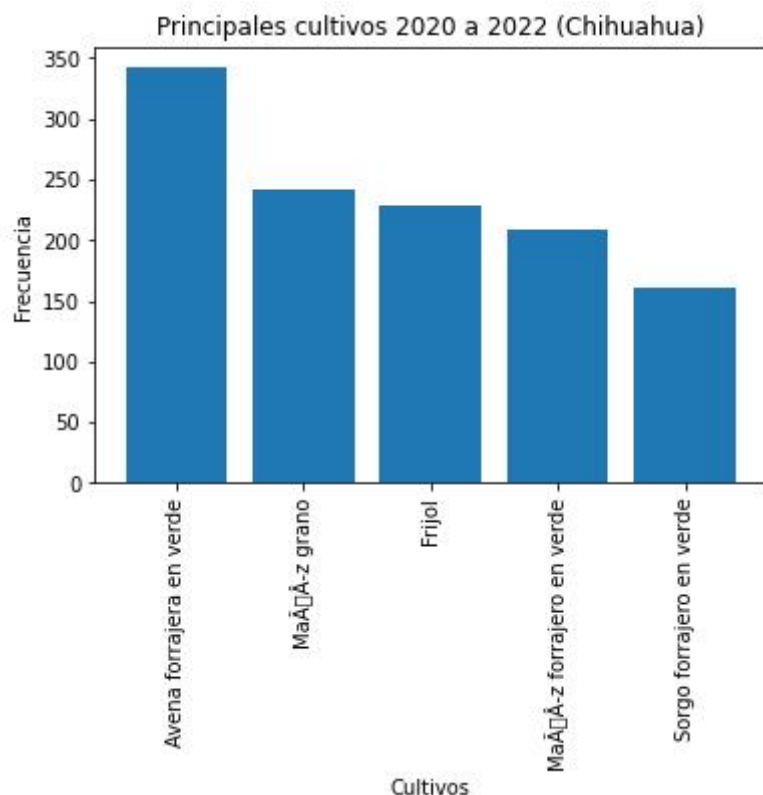
De 2020 a 2022 se produjeron 326 tipos de cultivos diferentes, de los cuales los principales fueron los siguientes:



Se logra concluir que nacionalmente se espera que los 5 principales cultivos sean el grano de maíz, el frijol, la avena forrajera en verde, el jitomate y el grano de sorgo. Además se espera una frecuencia de producción de entre 4500 y 10000 veces a lo largo de todos los municipios, tipos de riego, estaciones, etc. Para estos 5 principales cultivos.

Estatat:

De 2020 a 2022 se produjeron 65 tipos de cultivos diferentes en el estado de Chihuahua, de los cuales los principales fueron los siguientes:



(El mismo procedimiento puede ser replicado para obtener la información municipal y sobre cualquier dato relevante; en el caso de valores de toneladas, precios, etc. Se sumarían las cantidades de cada cultivo para obtener la información relevante. Además se puede caracterizar a cada estado como parte de una región y realizar el mismo procedimiento para encontrar valores relevantes solo que ampliando la condición a la región asignada y de interés).