

La pregunta acá fue “¿Cuáles son las causas de los incendios en cada zona?”.

Primero usé pandas para hacer dos groupby/sum por provincia: uno de foco y otro de superficie afectada. Exporté los resultados y los ploteé sobre mapas con [datawrapper](#).

En un mapa usé una escala de color blanco->rojo para representar cantidad de focos en cada provincia. En el otro mapa usé el tamaño de los círculos para representar superficie afectada, y en ese mismo mapa la escala de color blanco->rojo para representar cantidad de focos. Usé color rojo porque parecía natural para representar fuego, y porque después de mirar varios gráficos sobre los incendios de Australia hechos con bolitas rojas parecía ser la costumbre. (Ahí me habría gustado representar los focos con cantidades de bolitas en lugar de intensidad de color, pero no tenía esa opción en la herramienta online de datawrapper y no tuve tiempo para hacerlo a mano.)

En los dos mapas se puede ver que hay por lo menos 4 clusters: norte, este, centro, suroeste. A grandes rasgos, en norte/este/centro parece haber menos focos y más superficie afectada. En suroeste, más focos y menos superficie.

Para tratar de entender las causas, bajé otro [dataset](#) oficial con causas de incendios. Con pandas hice groupby/sum por provincia, para cada causa (natural, intencional o negligencia). Representé los resultados con stacked bars para mostrar la proporción de cada causa en cada provincia, y después agregué líneas verticales para separar las provincias en zonas (norte, este, centro, suroeste).

Como el gráfico de causas muestra diferencias entre el sur (más focos intencionales/por negligencia) y el centro (más focos naturales), lo puse encima del mapa que muestra los focos.

En el otro mapa quise ver las causas naturales. Bajé otro [dataset](#) oficial con mediciones del clima y filtré por los años del dataset del tp. Tomé los registros de estaciones de observación de las provincias del tp, groupby/sum, separé en zonas, y representé con scatterplots de líneas la variación por mes de lluvia y temperaturas máximas. Las líneas de cada subplot intentan mostrar mediciones parecidas, para correlacionar las 4 zonas con esos dos factores climáticos.

Aspectos que quedaron afuera:

- Como agrupé por provincia, no analicé los departamentos; los incendios parecen estar puntualmente localizados en ciertos lugares.
- Como agrupé por todo el período 2011-2016, no analicé las variaciones en distintos años. Hice eso porque la similitud entre las provincias de cada zona parece mantenerse a lo largo de los años, a juzgar por este copado [visualizador oficial](#), pero en cada año seguramente hay particularidades interesantes.
- Discriminar si el incendio ocurre en un parque nacional o en terreno privado.
- Hay otros factores climáticos que [influyen](#): nubosidad, humedad relativa, viento.