

# Applied Competitive Lab in Data Science

## Exercise 3

Deadline: 24.11.2022 at 23:55

### 0 Dataset Info

Your grade would be made from the top 3 questions you've answered. Context: This data publication contains a spatial database of wildfires that occurred in the United States from 1992 to 2015. It is the third update of a publication originally generated to support the national Fire Program Analysis (FPA) system. The wildfire records were acquired from the reporting systems of federal, state, and local fire organizations. The following core data elements were required for records to be included in this data publication: discovery date, final fire size, and a point location at least as precise as Public Land Survey System (PLSS) section (1-square mile grid). The data were transformed to conform, when possible, to the data standards of the National Wildfire Coordinating Group (NWCG). Basic error-checking was performed and redundant records were identified and removed, to the degree possible. The resulting product, referred to as the Fire Program Analysis fire-occurrence database (FPA FOD), includes 1.88 million geo-referenced wildfire records, representing a total of 140 million acres burned during the 24-year period.

### 1 Feature Information

- **FOD\_ID** = Global unique identifier.
- **FPA\_ID** = Unique identifier that contains information necessary to track back to the original record in the source dataset.
- **SOURCESYSTEMTYPE** = Type of source database or system that the record was drawn from (federal, nonfederal, or interagency).
- **SOURCESYSTEM** = Name of or other identifier for source database or system that the record was drawn from. See Table 1 in Short (2014), or .pdf, for a list of sources and their identifier.
- **NWCGREPORTINGAGENCY** = Active National Wildlife Coordinating Group (NWCG) Unit Identifier for the agency preparing the fire report (BIA = Bureau of Indian Affairs, BLM = Bureau of Land Management,

BOR = Bureau of Reclamation, DOD = Department of Defense, DOE = Department of Energy, FS = Forest Service, FWS = Fish and Wildlife Service, IA = Interagency Organization, NPS = National Park Service, ST/C&L = State, County, or Local Organization, and TRIBE = Tribal Organization).

- **NWCGREPORTINGUNIT\_ID** = Active NWCG Unit Identifier for the unit preparing the fire report.
- **NWCGREPORTINGUNIT\_NAME** = Active NWCG Unit Name for the unit preparing the fire report. **SOURCEREPORTINGUNIT** = Code for the agency unit preparing the fire report, based on code/name in the source dataset.
- **SOURCEREPORTINGUNIT\_NAME** = Name of reporting agency unit preparing the fire report, based on code/name in the source dataset.
- **LOCALFIREREPORT\_ID** = Number or code that uniquely identifies an incident report for a particular reporting unit and a particular calendar year.
- **LOCALINCIDENTID** = Number or code that uniquely identifies an incident for a particular local fire management organization within a particular calendar year.
- **FIRE.CODE** = Code used within the interagency wildland fire community to track and compile cost information for emergency fire suppression (<https://www.firecode.gov/>).
- **FIRE\_NAME** = Name of the incident, from the fire report (primary) or ICS-209 report (secondary).
- **ICS209INCIDENT\_NUMBER** = Incident (event) identifier, from the ICS-209 report.
- **ICS209NAME** = Name of the incident, from the ICS-209 report.
- **MTBS\_ID** = Incident identifier, from the MTBS perimeter dataset.
- **MTBSFIRENAME** = Name of the incident, from the MTBS perimeter dataset.
- **COMPLEX\_NAME** = Name of the complex under which the fire was ultimately managed, when discernible.
- **FIRE\_YEAR** = Calendar year in which the fire was discovered or confirmed to exist.
- **DISCOVERY\_DATE** = Date on which the fire was discovered or confirmed to exist.

- **DISCOVERY\_DOY** = Day of year on which the fire was discovered or confirmed to exist.
- **DISCOVERY\_TIME** = Time of day that the fire was discovered or confirmed to exist.
- **STATCAUSECODE** = Code for the (statistical) cause of the fire.
- **STATCAUSEDESCR** = Description of the (statistical) cause of the fire.
- **CONT\_DATE** = Date on which the fire was declared contained or otherwise controlled (mm/dd/yyyy where mm=month, dd=day, and yyyy=year).
- **CONT\_DOY** = Day of year on which the fire was declared contained or otherwise controlled.
- **CONT\_TIME** = Time of day that the fire was declared contained or otherwise controlled (hhmm where hh=hour, mm=minutes).
- **FIRE\_SIZE** = Estimate of acres within the final perimeter of the fire.
- **FIRESIZECLASS** = Code for fire size based on the number of acres within the final fire perimeter expenditures (A=greater than 0 but less than or equal to 0.25 acres, B=0.26-9.9 acres, C=10.0-99.9 acres, D=100-299 acres, E=300 to 999 acres, F=1000 to 4999 acres, and G=5000+ acres).
- **LATITUDE** = Latitude (NAD83) for point location of the fire (decimal degrees).
- **LONGITUDE** = Longitude (NAD83) for point location of the fire (decimal degrees).
- **OWNER\_CODE** = Code for primary owner or entity responsible for managing the land at the point of origin of the fire at the time of the incident.
- **OWNER\_DESCR** = Name of primary owner or entity responsible for managing the land at the point of origin of the fire at the time of the incident.
- **STATE** = Two-letter alphabetic code for the state in which the fire burned (or originated), based on the nominal designation in the fire report.
- **COUNTY** = County, or equivalent, in which the fire burned (or originated), based on nominal designation in the fire report.
- **FIPS\_CODE** = Three-digit code from the Federal Information Process Standards (FIPS) publication 6-4 for representation of counties and equivalent entities.
- **FIPS\_NAME** = County name from the FIPS publication 6-4 for representation of counties and equivalent entities.

- NWCGUnitIDActive20170109: Look-up table containing all NWCG identifiers for agency units that were active (i.e., valid) as of 9 January 2017, when the list was downloaded from [https://www.nifc.blm.gov/unit\\_id/Publish.html](https://www.nifc.blm.gov/unit_id/Publish.html) and used as the source of values available to populate the following fields in the Fires table: NWCGREPORTINGAGENCY, NWCGREPORTINGUNITID, and NWCGREPORTINGUNITNAME.
- UnitId = NWCG Unit ID.
- GeographicArea = Two-letter code for the geographic area in which the unit is located (NA=National, IN=International, AK=Alaska, CA=California, EA=Eastern Area, GB=Great Basin, NR=Northern Rockies, NW=Northwest, RM=Rocky Mountain, SA=Southern Area, and SW=Southwest).
- Gacc = Seven or eight-letter code for the Geographic Area Coordination Center in which the unit is located or primarily affiliated with.
- WildlandRole = Role of the unit within the wildland fire community.
- UnitType = Type of unit (e.g., federal, state, local).
- Department = Department (or state/territory) to which the unit belongs.
- Agency = Agency or bureau to which the unit belongs.
- Parent = Agency subgroup to which the unit belongs (A concatenation of State and Unit from this report - [https://www.nifc.blm.gov/unit\\_id/publish/UnitIdReport.rtf](https://www.nifc.blm.gov/unit_id/publish/UnitIdReport.rtf)).
- Country = Country in which the unit is located (e.g. US = United States).
- State = Two-letter code for the state in which the unit is located (or primarily affiliated).
- Code = Unit code (follows state code to create UnitId).
- Name = Unit name.