Database Variable Breakdown

The provided variables contain a lot of excess information. Below we go through all of the provided variables. The list is provided on the Kaggle site that provides the database, (1). There is an excess of variables that are not predictors but identifying indicators for the authorities to file the proper paperwork regarding the fire. These variables are all crossed out below. We have the variables related to the response, cause of the fire, highlighted in blue and the potential predictors highlighted in yellow. It would be very problematic to fit a model containing all of these potential predictors because there are a lot of redundancies. One obvious example is fire size and fire size class are measuring the same thing in two different ways. Now that we have narrowed down which variables may be useful we can move forward to predictor selection being careful not to include two variables that are actually measuring the same thing.

This dataset is an SQLite database that contains the following information:

- Fires: Table including wildfire data for the period of 1992-2015 compiled from US federal, state, and local reporting systems.
- FOD_ID = Global unique identifier.
- FPA_ID = Unique identifier that contains information necessary to track back to the original record in the source dataset.
- SOURCE_SYSTEM_TYPE = Type of source database or system that the record was drawn from (federal, nonfederal, or interagency).
- SOURCE_SYSTEM = Name of or other identifier for source database or system that the record was drawn from. See Table 1 in Short (2014), or \Supplements\FPA_FOD_source_list.pdf, for a list of sources and their identifier.
- NWCG_REPORTING_AGENCY = 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).
- NWCG_REPORTING_UNIT_ID = Active NWCG Unit Identifier for the unit preparing the fire report.
- NWCG_REPORTING_UNIT_NAME = Active NWCG Unit Name for the unit preparing the fire report.
- SOURCE_REPORTING_UNIT = Code for the agency unit preparing the fire report, based on code/name in the source dataset.
- SOURCE_REPORTING_UNIT_NAME = Name of reporting agency unit preparing the fire report, based on code/name in the source dataset.
- LOCAL_FIRE_REPORT_ID = Number or code that uniquely identifies an incident report for a particular reporting unit and a particular calendar year.

- LOCAL_INCIDENT_ID = 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).
- ICS_209_INCIDENT_NUMBER = Incident (event) identifier, from the ICS-209 report.
- ICS_209_NAME = Name of the incident, from the ICS-209 report.
- MTBS_ID = Incident identifier, from the MTBS perimeter dataset.
- MTBS_FIRE_NAME = 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.
- STAT_CAUSE_CODE = Code for the (statistical) cause of the fire.
- STAT_CAUSE_DESCR = 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.
- FIRE_SIZE_CLASS = 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.

- NWCG_UnitIDActive_20170109: 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 and used as the source of values available to populate the following fields in the Fires table: NWCG_REPORTING_AGENCY, NWCG_REPORTING_UNIT_ID, and NWCG_REPORTING_UNIT_NAME.
- 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 (CAMBCIFC=Canadian Interagency Forest Fire Centre, USAKCC=Alaska Interagency Coordination Center, USCAONCC=Northern California Area Coordination Center, USCAOSCC=Southern California Coordination Center, USCORMCC=Rocky Mountain Area Coordination Center, USGASAC=Southern Area Coordination Center, USIDNIC=National Interagency Coordination Center, USMTNRC=Northern Rockies Coordination Center, USNMSWC=Southwest Area Coordination Center, USORNWC=Northwest Area Coordination Center, USUTGBC=Western Great Basin Coordination Center, USWIEACC=Eastern Area Coordination Center).
- 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 (AK=Alaska, AL=Alabama, AR=Arkansas, AZ=Arizona, CA=California, CO=Colorado, CT=Connecticut, DE=Delaware, DHS=Department of Homeland Security, DOC= Department of Commerce, DOD=Department of Defense, DOE=Department of Energy, DOI= Department of Interior, DOL=Department of Labor, FL=Florida, GA=Georgia, IA=Iowa, IA/GC=Non-Departmental Agencies, ID=Idaho, IL=Illinois, IN=Indiana, KS=Kansas, KY=Kentucky, LA=Louisiana, MA=Massachusetts, MD=Maryland, ME=Maine, MI=Michigan, MN=Minnesota, MO=Missouri, MS=Mississippi, MT=Montana, NC=North Carolina, NE=Nebraska, NG=Non-Government, NH=New Hampshire, NJ=New Jersey, NM=New Mexico, NV=Nevada, NY=New York, OH=Ohio, OK=Oklahoma, OR=Oregon, PA=Pennsylvania, PR=Puerto Rico, RI=Rhode Island, SC=South Carolina, SD=South Dakota, ST/L=State or Local Government, TN=Tennessee, Tribe=Tribe, TX=Texas, USDA=Department of Agriculture, UT=Utah, VA=Virginia, VI=U. S. Virgin Islands, VT=Vermont, WA=Washington, WI=Wisconsin, WV=West Virginia, WY=Wyoming).
- Agency = Agency or bureau to which the unit belongs (AG=Air Guard, ANC=Alaska Native Corporation, BIA=Bureau of Indian Affairs, BLM=Bureau of Land Management, BOEM=Bureau of Ocean Energy Management, BOR=Bureau of Reclamation, BSEE=Bureau of Safety and Environmental Enforcement, C&L=County & Local, CDF=California Department of Forestry & Fire Protection, DC=Department of Corrections, DFE=Division of Forest Environment, DFF=Division of Forestry Fire & State Lands, DFL=Division of Forests and Land, DFR=Division of Forest Resources, DL=Department of Lands, DNR=Department of Natural Resources,

DNRC=Department of Natural Resources and Conservation, DNRF=Department of Natural Resources Forest Service, DOA=Department of Agriculture, DOC=Department of Conservation, DOE=Department of Energy, DOF=Department of Forestry, DVF=Division of Forestry, DWF=Division of Wildland Fire, EPA=Environmental Protection Agency, FC=Forestry Commission, FEMA=Federal Emergency Management Agency, FFC=Bureau of Forest Fire Control, FFP=Forest Fire Protection, FFS=Forest Fire Service, FR=Forest Rangers, FS=Forest Service, FWS=Fish & Wildlife Service, HQ=Headquarters, JC=Job Corps, NBC=National Business Center, NG=National Guard, NNSA=National Nuclear Security Administration, NPS=National Park Service, NWS=National Weather Service, OES=Office of Emergency Services, PRI=Private, SF=State Forestry, SFS=State Forest Service, SP=State Parks, TNC=The Nature Conservancy, USA=United States Army, USACE=United States Army Corps of Engineers, USAF=United States Air Force, USGS=United States Geological Survey, USN=United States Navy).

- 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).
- 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.

Sources:

1. Tatman, Rachael. "1.88 Million US Wildfires | Kaggle." Countries of the World | Kaggle, 13 Sept. 2017, www.kaggle.com/rtatman/188-million-us-wildfires.