Project Proposal
California Wildfires Analysis

DSE241 Sergey Gurvich

Source: https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcSuMwWPAUHqfUUxuGHag9YPKDWZvlx0fNcCbrtvH3nlHqm1IIR&s

The Dataset: Historical California Wildfire Data

Source: https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcQxmUVirXtPwk6P3S4224iIFQDCm62Fiihfntb8AB4HiDHRpUs&s

The California Department of Forestry and Fire Protection (CAL FIRE) maintains historical data about wildfires in California
The dataset contains information on California fires dating back to 1878, providing the following summary information for each identified fire:

- YEAR
- STATE
- AGENCY
- UNIT_ID
- FIRE_NAME
- INC_NUM (incident number)
- ALARM_DATE
- CONT_DATE (containment date)
- CAUSE
- COMMENTS
- REPORT_AC (acres burned)
- GIS_ACRES (acres burned: GIS data)
- C_METHOD
- OBJECTIVE

YEAR	UNIT_ID	FIRE_NAME	ALARM_DATE	CONT_DATE	CAUSE	REPORT_AC	GIS_ACRES	C_METHOD	OBJECTIVE
2008	Shasta-Trinity National Forest	SHU LIGHTNING-KIRKMAN	6/20/2008	<Null>	1 - Lightning	2054	2053.58	2 - GPS Air	Suppression (Wildfire)
2008	Sequoia National Forest	MOSES	10/15/2008	10/31/2008	1 - Lightning	105	231.5684	2 - GPS Air	Resource Benefit (WFU)
2008	Tehama - Glenn CAL FIRE	MILL	6/21/2008	6/29/2008	1 - Lightning	13580	13511.76	8 - Unknown	Suppression (Wildfire)
2008	Plumas National Forest	FOUR MILE	6/21/2008	<Null>	1 - Lightning	<Null>	788.7921	8 - Unknown	Suppression (Wildfire)
2008	San Joaquin River National Wildlife Refuge	RIVER	6/24/2008	6/25/2008	9 - Miscellaneous	580	573.1215	8 - Unknown	<Null>
2008	Lava Beds National Monument	JACK	8/17/2008	8/31/2008	1 - Lightning	5397	6067.5	7 - Mixed Collection Methods	Suppression (Wildfire)
2008	Sequoia - Kings Canyon NP	HIDDEN	9/10/2008	10/8/2008	1 - Lightning	3685	3685.978	7 - Mixed Collection Methods	Suppression (Wildfire)
2008	Sequoia - Kings Canyon NP	TEHIPITE	7/19/2008	12/1/2008	1 - Lightning	4140	11648.09	3 - Infrared	Resource Benefit (WFU)
2008	CA Desert District - BLM	MONTGOMERY	4/7/2008	4/7/2008	4 - Campfire	290	289.9371	1 - GPS Ground	Suppression (Wildfire)
2009	CRA	DEER	7/6/2009	7/21/2009	14 - Unknown / Unidentified	<Null>	330.1603	8 - Unknown	Suppression (Wildfire)
2009	Vandenberg Air Force Base	HIGHWAY	9/30/2009	10/3/2009	11 - Powerline	562.127013	562.13	8 - Unknown	Suppression (Wildfire)
2009	Angeles National Forest	TUJUNGA	7/5/2009	7/9/2009	1 - Lightning	180	210.5635	7 - Mixed Collection Methods	Suppression (Wildfire)
2009	Angeles National Forest	OSITO	7/13/2009	7/13/2009	2 - Equipment Use	300	332.0352	2 - GPS Air	Suppression (Wildfire)
2009	Angeles National Forest	MORRIS	8/25/2009	11/30/2009	7 - Arson	2163	2236.504	7 - Mixed Collection Methods	Suppression (Wildfire)
2009	Angeles National Forest	STATION	8/26/2009	9/22/2009	7 - Arson	160371	160833.1	7 - Mixed Collection Methods	Suppression (Wildfire)
2009	San Bernardino National Forest	MCKINLEY	6/3/2009	6/3/2009	1 - Lightning	150	166.3999	6 - Hand Drawn	Suppression (Wildfire)
2009	San Bernardino National Forest	ELM	7/7/2009	7/7/2009	9 - Miscellaneous	350	348.3582	6 - Hand Drawn	Suppression (Wildfire)
2009	San Bernardino National Forest	COTTONWOOD	8/27/2009	8/31/2009	9 - Miscellaneous	2409	2410.618	7 - Mixed Collection Methods	Suppression (Wildfire)
2009	San Bernardino National Forest	SHEEP	10/3/2009	10/24/2009	9 - Miscellaneous	7128	7237.873	7 - Mixed Collection Methods	Suppression (Wildfire)
2009	San Bernardino CAL FIRE	CRAFTON	9/23/2009	9/27/2009	14 - Unknown / Unidentified	347	348.1416	1 - GPS Ground	Suppression (Wildfire)
2009	San Bernardino CAL FIRE	FORT	2/5/2009	2/7/2009	18 - Escaped Prescribed Burn	945	943.448	1 - GPS Ground	Suppression (Wildfire)
2009	San Bernardino CAL FIRE	OAK GLEN 3	8/30/2009	9/10/2009	7 - Arson	1159	983.2415	2 - GPS Air	Suppression (Wildfire)
2009	San Bernardino CAL FIRE	PENDLETON	8/31/2009	9/3/2009	7 - Arson	860	838.701	2 - GPS Air	Suppression (Wildfire)
2009	Monterey - San Benito CAL FIRE	BRYSON	8/26/2009	8/28/2009	15 - Structure	<Null>	2256.828	7 - Mixed Collection Methods	Suppression (Wildfire)
2009	Monterey - San Benito CAL FIRE	GLORIA	8/28/2009	9/8/2009	9 - Miscellaneous	<Null>	6435.952	2 - GPS Air	Suppression (Wildfire)
2009	CA Desert District - BLM	LAMB	7/11/2009	7/11/2009	10 - Vehicle	288	288.1363	7 - Mixed Collection Methods	Suppression (Wildfire)
2009	CA Desert District - BLM	MIDWAY	4/17/2009	<Null>	2 - Equipment Use	112	111.6097	1 - GPS Ground	Suppression (Wildfire)
2009	San Mateo - Santa Cruz CAL FIRE	LOCKHEED	8/12/2009	8/23/2009	14 - Unknown / Unidentified	7817	7783.064	1 - GPS Ground	Suppression (Wildfire)
2009	San Mateo - Santa Cruz CAL FIRE	LOMA	10/25/2009	10/27/2009	14 - Unknown / Unidentified	650	669.3967	1 - GPS Ground	Suppression (Wildfire)
2009	Hoopa Valley Tribe	MILL CREEK 4	10/7/2009	11/18/2009	9 - Miscellaneous	3689	2831.125	8 - Unknown	Suppression (Wildfire)
2009	Humboldt - Del Norte CAL FIRE	PRATT	9/11/2009	9/11/2009	2 - Equipment Use	140	141.965	2 - GPS Air	Suppression (Wildfire)

Dataset Source: https://cvw.cac.cornell.edu/PyDataSci1/wildfires

Task 1: Wildfires Geography Visualization

Visualization Purpose:	Presentation
Problem:	Identify the most affected geographical areas in California
Data Transformation/Augmentation:	Get geographical coordinates of fire locations by using geofencing API (passing Unit ID as address)
Use Case Scenarios:	Can be used for presentation to stakeholders for any wildfire related project
Idiom:	Bubble Map
Interaction:	Zoom, Pan, Mouse Hovering to see details
Implementation Tools:	Python / Jupyter Lab / Pandas / Plotly library
Stakeholders:	Any fire management agency, general public
Stretch Goal:	Add interaction to switch years / timeline

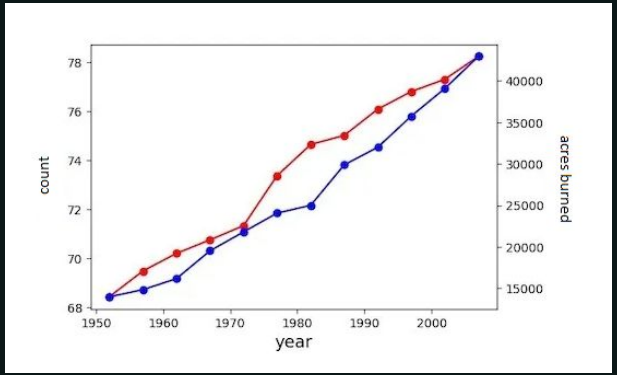
Sketch: California WildFires Geography

Souce: https://geology.com/state-map/california.shtml

Task 2: Wildfires Trends

Visualization Purpose:	Hypothesis Confirmation
Problem:	The wildfire activity has been increased and intensified over time due to climate change, population growth, undertreatment of forestry areas
Data Transformation/Augmentation:	Grouping
Use Case Scenarios:	Can be used for presentation to stakeholders for climate related projects
Idiom:	Line Plot (Dual-Axes)
Interaction:	Switch between Count and Area Burned views, Mouse Hovering to see details
Implementation Tools:	Python / Jupyter Lab / Pandas / Plotly / Bokeh / Seaborn / Matplotlib lib.
Stakeholders:	Any climate, fire management agency, general public

Sketch: dual-axes line plot

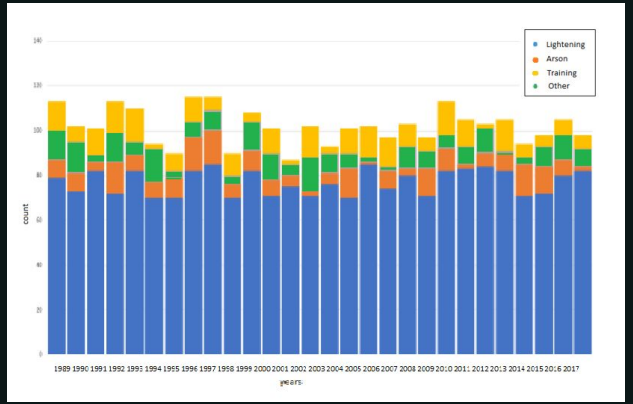


Souce: https://cmdlinetips.com/2019/10/how-to-make-a-plot-with-two-different-y-axis-in-python-with-matplotlib/

Task 3: Causes Of Wildfires Onset Analysis

Visualization Purpose:	Data Exploration
Problem:	Identify the main cause of wildfires: human or nature related?
Data Transformation/Augmentation:	Grouping
Use Case Scenarios:	Can be used for presentation to the fire management agency
Idiom:	Stacked Bar Chart
Interaction:	Filtering
Implementation Tools:	Python / Jupyter Lab / Pandas / Plotly / Bokeh / Seaborn / Matplotlib lib.
Stakeholders:	Any fire management agency, general public

Sketch: Stacked bar chart

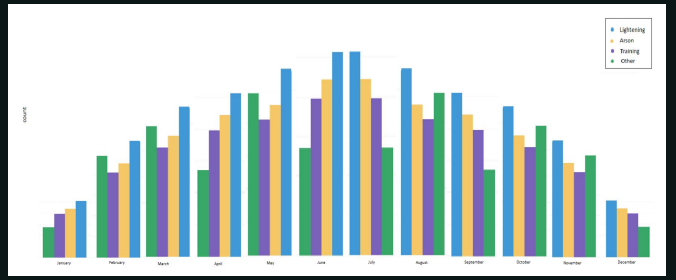


Source: https://community.looker.com/dashboards-looks-7/clustered-stacked-bar-chart-20876

Task 4: Seasonality Analysis

Visualization Purpose:	Hypothesis Confirmation
Problem:	When does the California wildfire season happen?
Data Transformation/Augmentation:	Grouping
Use Case Scenarios:	Can be used for presentation to stakeholders for fire related projects, planning
Idiom:	Grouped Bar Chart
Interaction:	Switch between Count and Area Burned views, Mouse Hovering to see details
Implementation Tools:	Python / Jupyter Lab / Pandas / Plotly / Bokeh / Seaborn / Matplotlib lib.
Stakeholders:	Any climate, fire management agency, general public

Sketch: Grouped Bar Chart

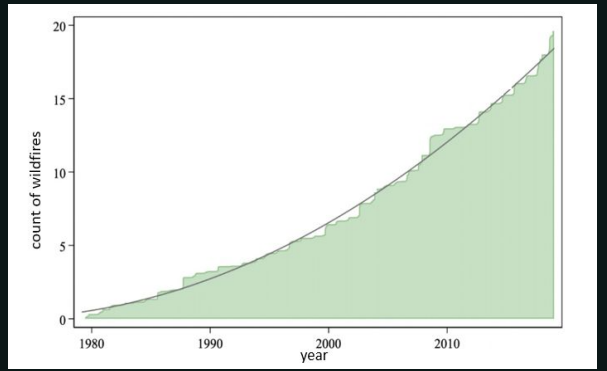


Souce: https://chartio.com/learn/charts/grouped-bar-chart-complete-guide/

Task 5: Fire Suppression Performance Analysis

Visualization Purpose:	Presentation
Problem:	Do modern suppression techniques increase performance - decrease mean duration of wildfires?
Data Transformation/Augmentation:	New Column: Calculate the duration of each fire. Grouping by years: mean fire duration.
Use Case Scenarios:	Can be used for presentation to fire agency supervision
Idiom:	Area Plot
Interaction:	Mouse Hovering to see details
Implementation Tools:	Python / Jupyter Lab / Pandas / Plotly / Bokeh / Seaborn / Matplotlib lib.
Stakeholders:	Any fire management agency, general public

Sketch: area plot



Souce: https://wildfiretoday.com/2020/03/31/data-shows-the-worsening-trend-of-california-wildfires/

Source: https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcQ_kXIhGVBs1uPAuG0s7YI9R9xzWhvkbqn9emSibi5N-RD2vbA&s