VOLCANIC ACTIVITY

Nibedita Bal, Jared Hebda, Evelina Ramoskaite, John Aucar

NW2 DSC 465 Summer 2020 Final Project

Table of Contents

Introduction

Volcanic eruptions happen when lava and gas are discharged from a volcanic vent. The most common consequences of this are population movements as large numbers of people are often forced to flee the moving lava flow. The flow moves down at a very high speed like several hundred km per hour which makes some animals, cities, and lands extinct forever. Eruption boosts world temperature, greenhouse gases, and sub-seafloor methane. In this project, we will deal with the past some years of volcanic activities along with the tectonic plate boundaries movement. We have around 8 hypotheses about the eruption and they will be visualized to give users enough facts and information to get deeper insight. Some exploratory data analyses are done prior to starting to visualize the hypothesis.

Data Overview

The datasets are from The Smithsonian Institution and contain 5 files related to eruption events, sulfur, tree rings, and volcanoes. There are 57 features in total. The data comes from all around the world, with eruptions dating back to 11,345 BCE. Below are the datasets and the sources that come from.

The two datasets, Volcano Data, and Tectonic plate boundaries are subdivided with 6 different datasets, as described below:

a.Eruptions file : consists 11,179 items and 11 variables

b.Events file : consists 41,323 items and 8 variables

c.Sulfur file : consists 2,253 items and 3 variables

d.Tree Rings file : consists 2,253 items and 3 variables

e.Volcano file : consists 959 items and 14 variables

F.all file : consists 12,322 items and 3 variables

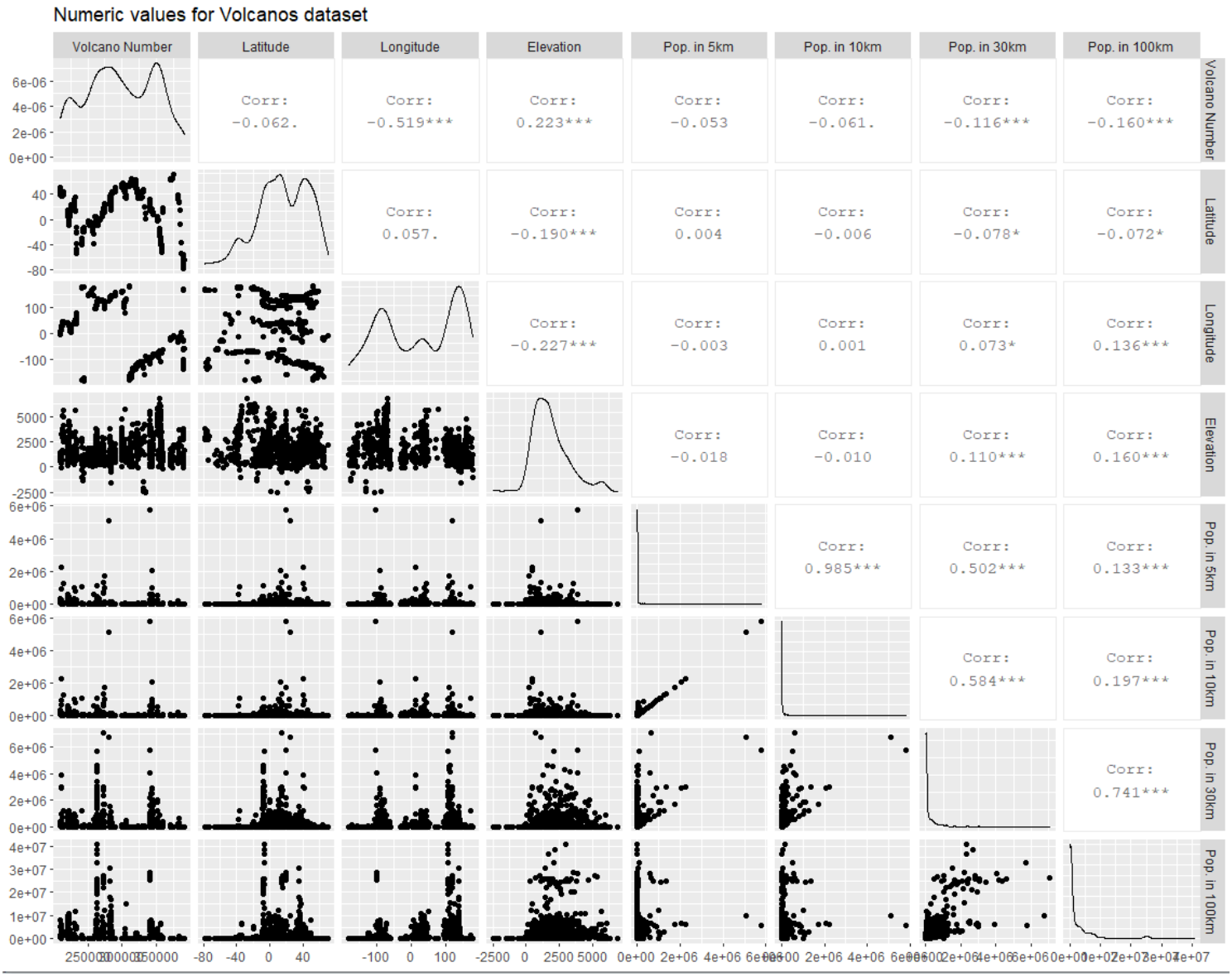
As we have 57 features in total, only important variables that are used in this project are mentioned below:

Field	**Description**
volcano_number	Every volcano has a unique number
volcano_name	Volcano name
primary_volcano_type	Type of volcanoes
last_eruption_year	For every volcano, the last year the volcano has erupted
country	Counties have volcanic eruptions
region	Regions of countries have eruptions
latitude	Latitude (low to mid to high)
longitude	Longitude
major_rock_1,2,3,4,5	Eruptions depend on some type of the rocks and here 5 types of rocks are discussed
minor_rock_1,2,3,4,5	Same as major rocks
Population_within_5 km,10 km,30 km,100 km	Population from the place of eruptions
year	Tree ring produced based on the year
n_tree	Tree ring
europe_temp_index	Temp is measured based on the eruption
plate	Tectonic plates
area_of_activity	Eruptions are active in some areas
Start_year, month, day	Start year, month and date of eruptions
End_year, month, day	The end time of eruptions
eruption_number	Every eruption has a unique number
Eruption_start_year	Start year of eruption

Exploratory Analysis

The dataset that we obtained from Kaggle has some missing values so the first step in our data exploration is to clean the dataset by imputing missing values or removing records entirely. After our dataset is cleaned, we will run summary statistics to help us understand how and why this data was collected. These summary statistics will include the distributions of numerical values and counts of categorical values. Missing values often occur in this dataset with eruptions that date back several thousand years. After we understand how the data is aggregated, we will begin to explore how the data interacts.

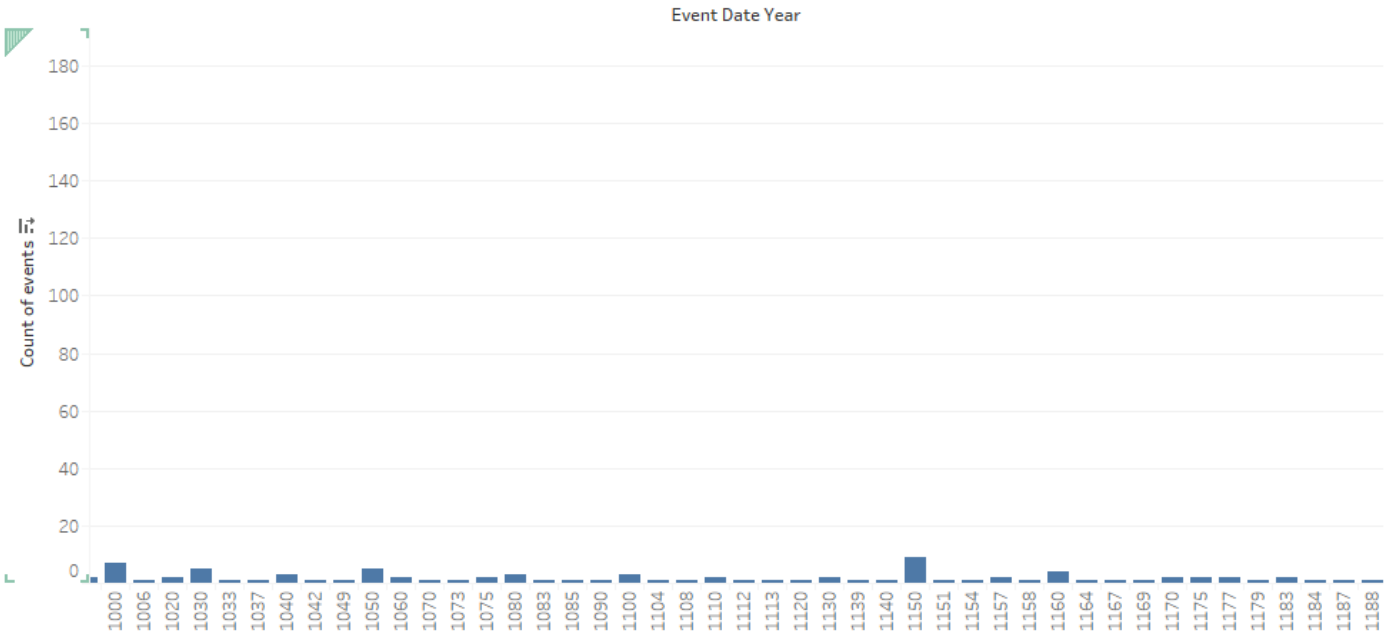
Next, we tested several interactions between variables to help gain insight into future eruptions and the effects of eruptions. Our data contains location information such as country, regions, latitude, and longitude which will help us visualize the results geographically. We also want to compare historical locations of eruptions with more current incidents to see if the geothermal activity has changed over time. Next, records will be separated categorically into the type of event, type of volcano, and other features to find patterns to determine the severity of the eruptions.



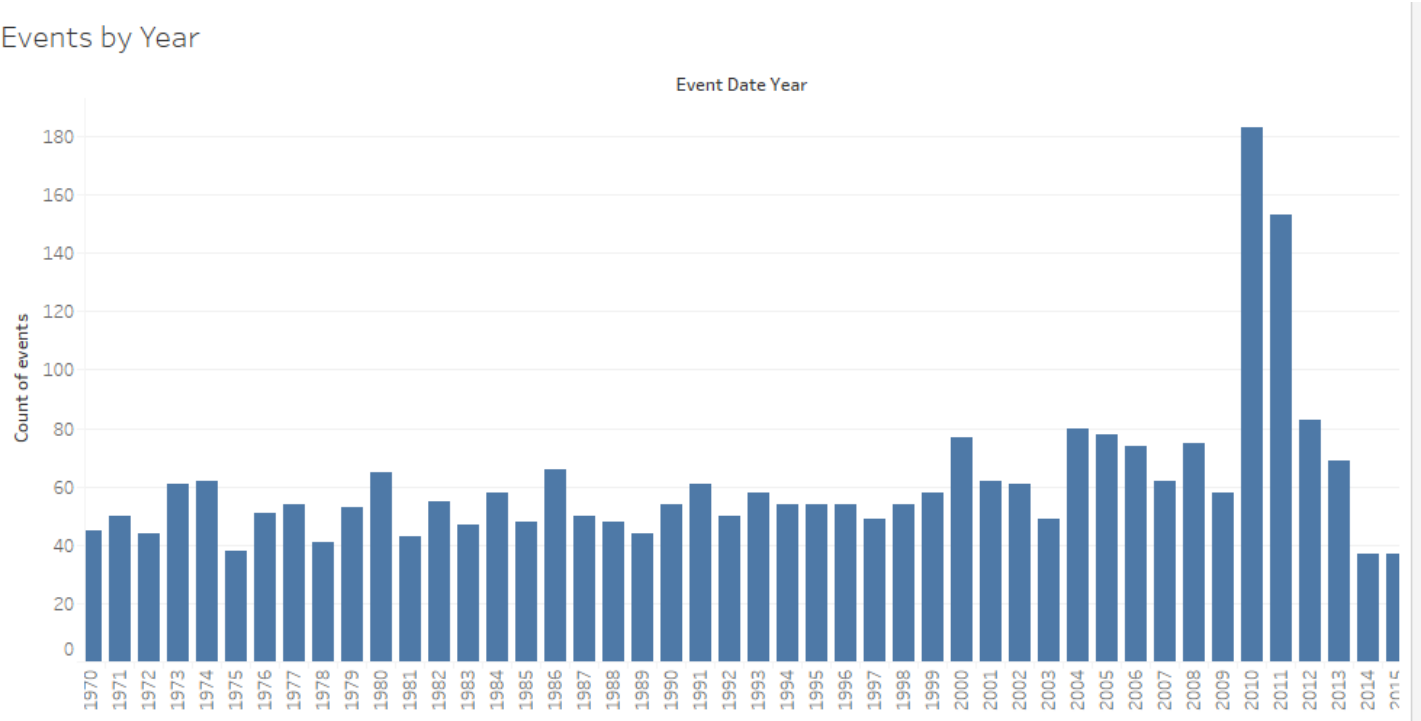
Events:

There are less than 20 events per year in this dataset until recent years. Events go back to -9650 and until 2020. From this, it is clear that event recording has improved over time and until recent days, certain technology wasn't available to measure some of these event types.

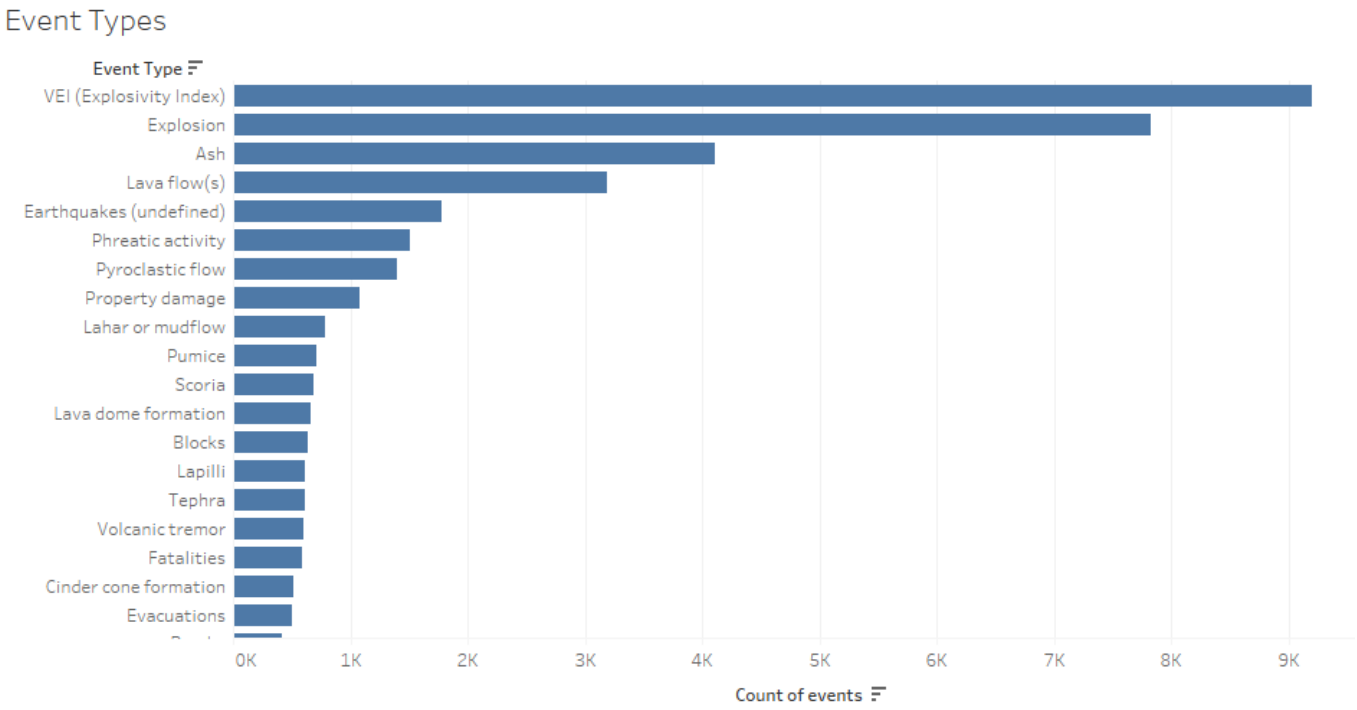
Events by Year



Events by Year



There are 56 distinct event types. The most common are VEI (Explosive Index), Explosion, Ash, Lava flow(s), and Earthquake (undefined). Since VEI (Explosive Index) didn't explain much about an event that occurred and rather was just a historical record of an eruption, we filtered it out of many of the visualizations.



31,315 event start years are not available out of 41,322.

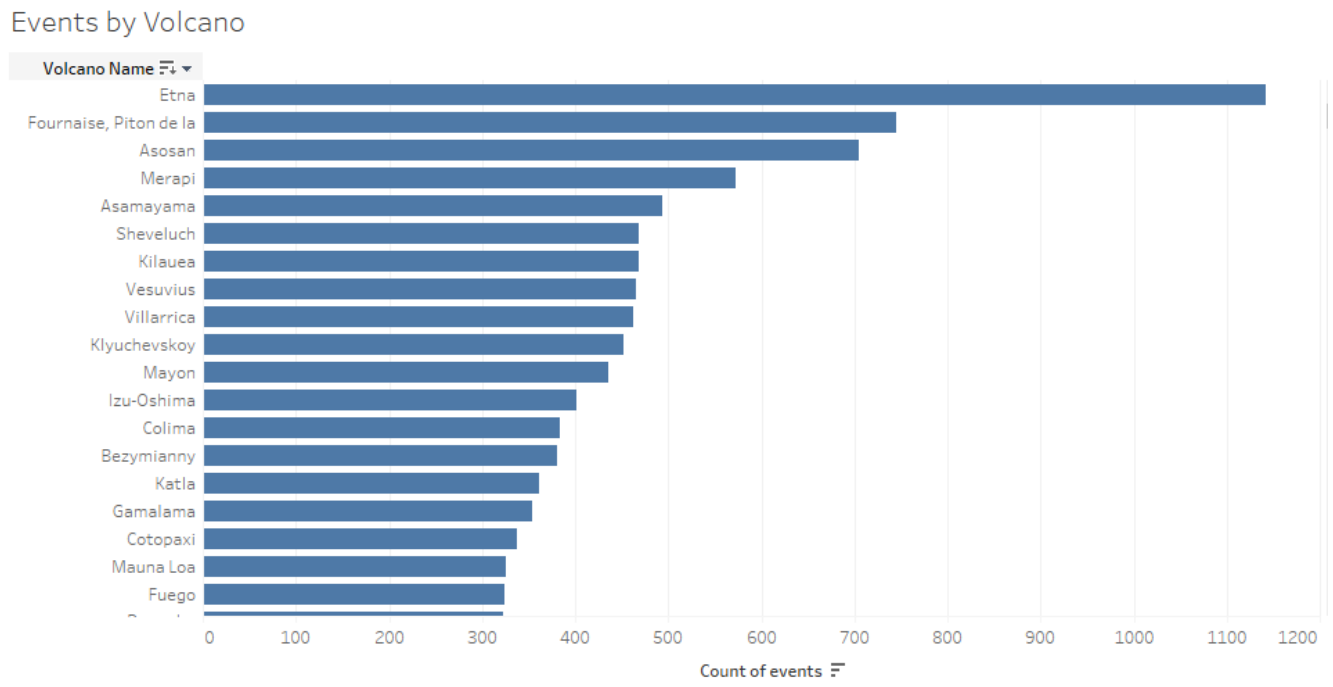
Several of the event types weren't being detected until recent years:

```
> sqldf("SELECT event_type, min(event_date_year)
+                 FROM events
+                 group by event_type
+                 order by min(event_date_year) desc
+
+                         ")
                         event_type min(event_date_year)
1                   Edifice destroyed                2014
2                  Loud audible noises               2013
3                           Degassing                2012
4                      Spine formation               2012
5                 Cinder cone formation              2011
6               Deformation (undefined)              2011
7                      Eruption cloud                2011
8                                Glow                2011
9                     Island formation               2011
10                     Lava fountains                2011
11                             Pumice                2011
12                    Volcanic tremor                2011
13                          Ash Plume                2010
14                            Ashfall                2010
15                             Blocks                2010
16                    Crater formation               2010
17                   Fissure formation               2010
18                  Incandescent ejecta              2010
19                  Lava dome formation              2010
20                           Lightning               2010
21 Partial collapse at end of eruption               2010
22                   Phreatic activity               2010
23                             Tephra                2010
24                              Bombs                2009
25                         Lava lake                  2006
26                        Observation                 2004
27                             Scoria                2004
28                    Thermal anomaly                2000
29                   Debris avalanches               1963
30                    Lahar or mudflow               1963
```

```
> sqldf("SELECT count(distinct volcano_number)
+                     from events
+                                 ")
  count(distinct volcano_number)
1                            907
```

There are 907 volcanoes with eruptions corresponding to 41,323 events in this file.

The following is a graph of events by volcano. This does not necessarily mean these volcanoes have more activity or events, but maybe they are more studied or analyzed than other volcanoes.



Plotting events with color offers some interesting insights when looking around the world, such as lava activity in Idaho. I would like to take this another step and have a visualization where you can filter on the type of event and see where in the world that event type happens. With the below image, it's hard to see event types since a volcano has several events associated with it. This led us to believe there could be some insights to plotting each volcano and filtering events.

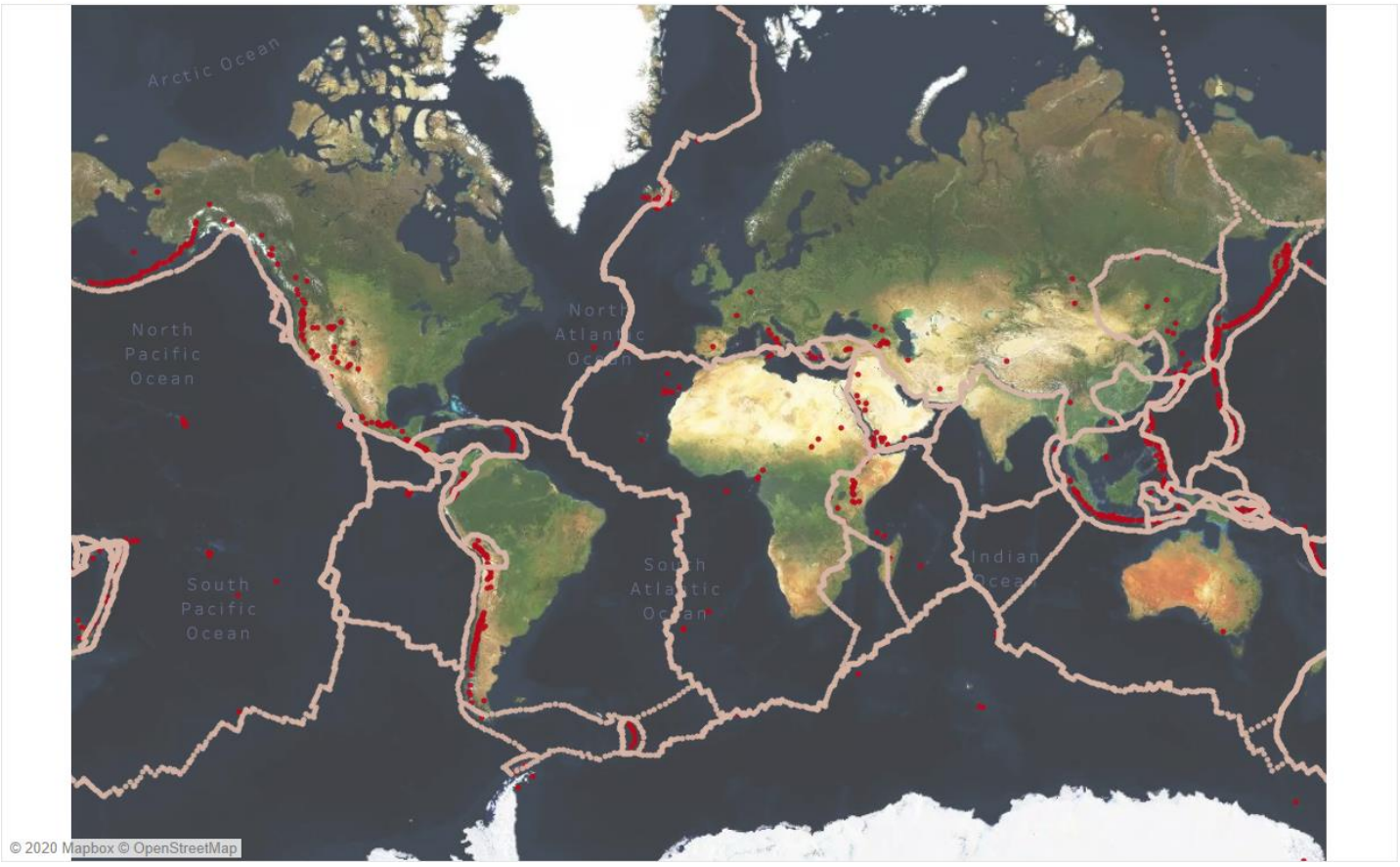
Hypothesis:

1. Volcanoes are likely aligned on fault lines:

Fault lines are the lines where the earth's crust plates strike one another. These lines are very likely to be placed where seismic activity occurs that means the shaking of the earth like volcanos. Therefore, fault lines are very dangerous to live and the fault lines can be observed by **map** plotted by plate variable data.

In our datasets_652326_1154075_all.csv, 3 variables latitude, longitude, and plate are considered to observe the fault lines in a world map.

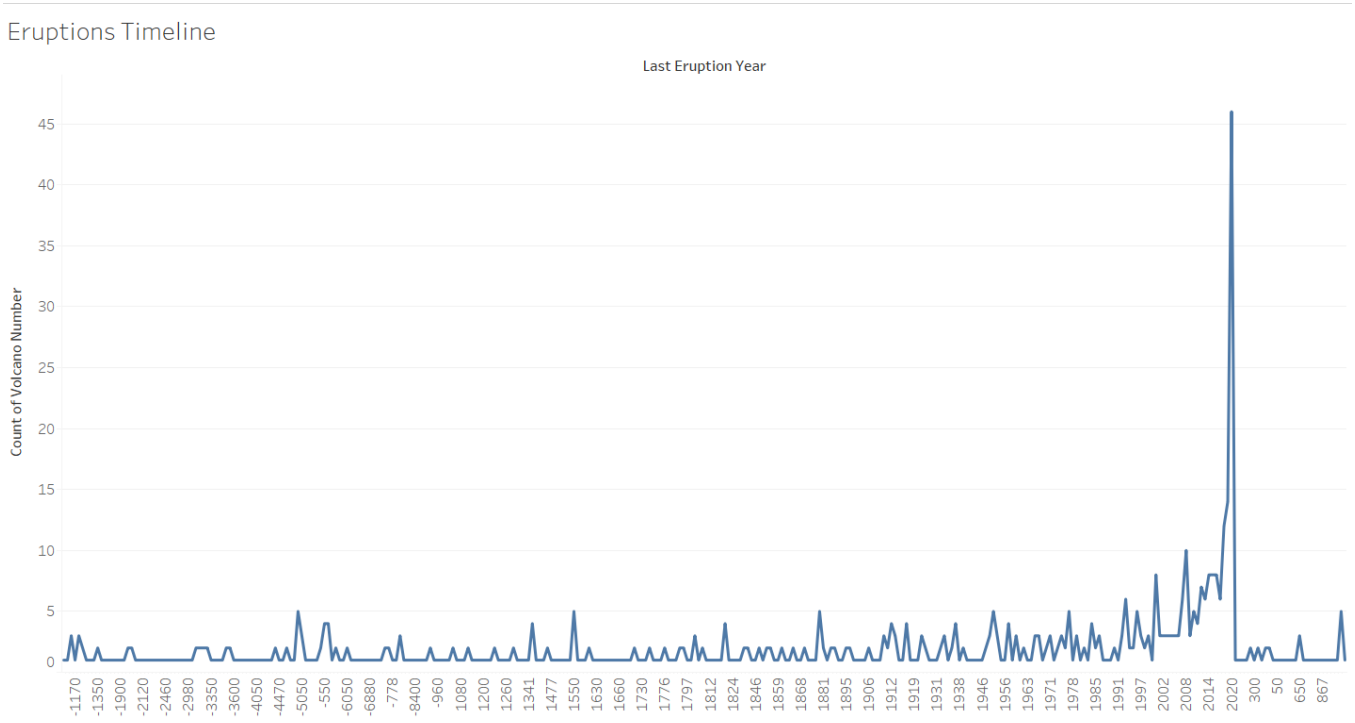
Map of Tectonic Plates and Volcanoes



2. Eruptions may occur in periodic increments.

According to The Smithsonian Institution of global volcanism program, there is no solid evidence of increased volcanic activity in past years. The last 200 years of eruption record with humans distributed over most of the globe and relatively efficient communications, would seem to be the well-suited to search for episodic trends. People living near the volcanoes observe the pattern well and the advanced communication technologies report these eruptions correctly. As a result, every small and large eruption that occurred in a remote area can be reported.

The periodic pattern can be well observed in a **time series graph or heatmap** where the graph will be plotted among variables last eruption years, region, and volcano number from the datasets_648337_1148296_volcano.

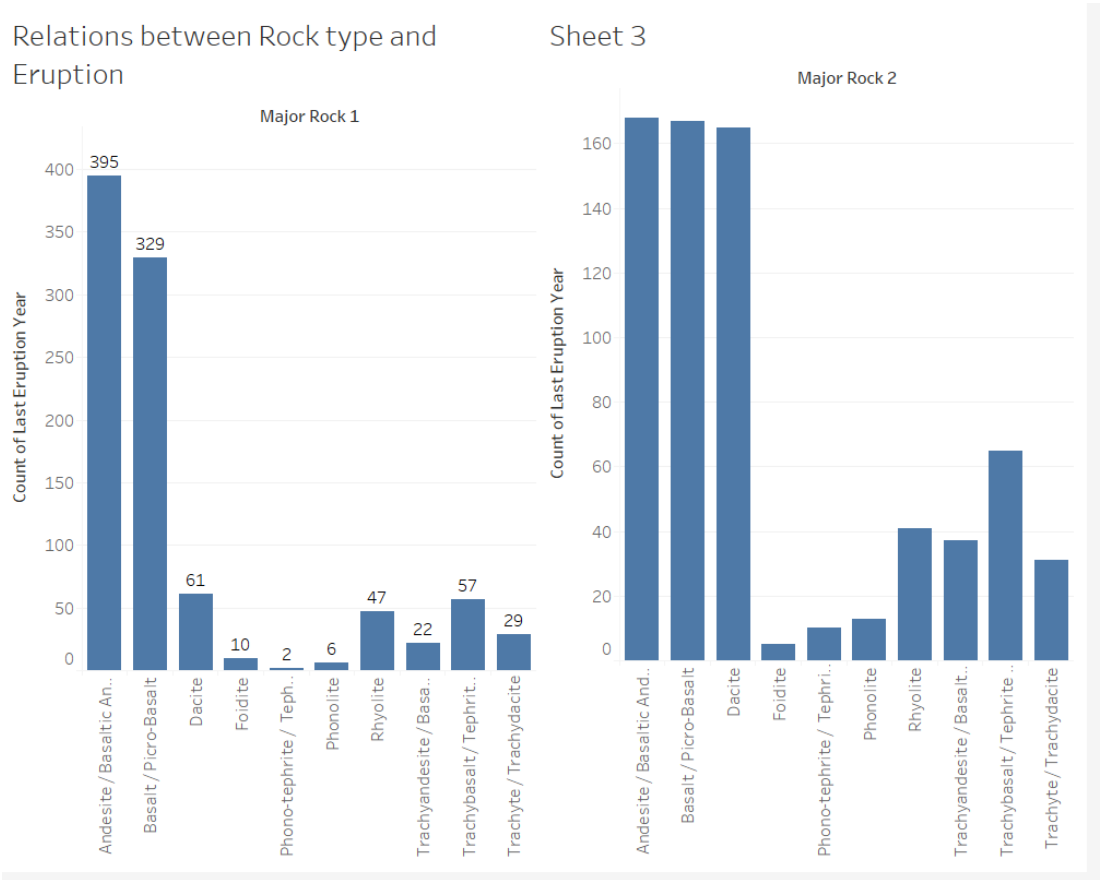


3. There may be a relationship between the rock type and frequency of eruption.

Volcanic rocks are formed from lava that flows on the surface of the Earth and then cools and becomes hard. The texture of an igneous rock depends on the size of the crystals in the rock which tells if a rock is volcanic or not. There may be a connection between the rock type and the occurrence of the eruption as Magma which comes out during the eruption is less dense than the surrounding rock which causes the eruption to rise. When magma reaches the surface it is then called lava and the eruptions of lava and ash produce volcanoes.

The datasets_648337_1148296_volcano consists of variables major rock 1,2,3,4,5 and minor rock 1,2,3,4,5 which can be plotted against the eruption year to show the relationships among them in a **bar plot**.

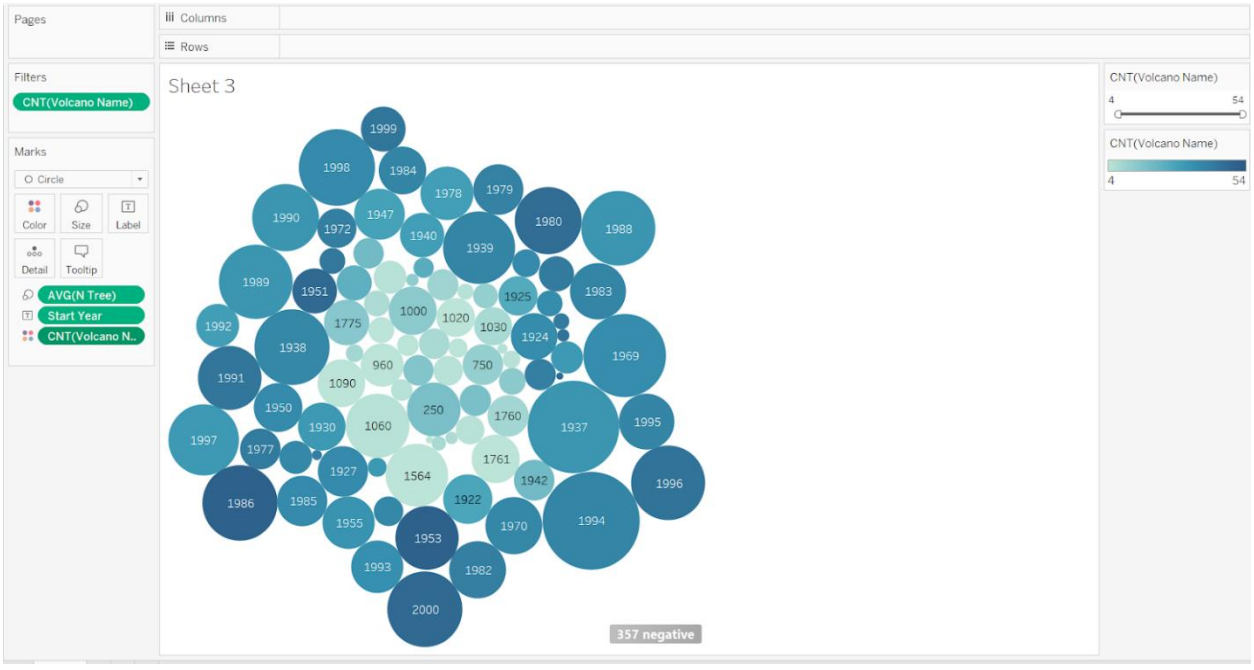
Major rocks 1,2,3,4,5:



4. Eruption frequency has an impact on the tree ring patterns of a given year.

There might be some relation between the tree ring patterns and the eruption frequency. Tree rings may be able to predict the eruption. The ring width may also be influenced by volcanic activity in the volcanic regions. knowing when and in what conditions a volcano erupted previously may help people to understand and predict when it could erupt again. However, climate affects the amount a tree grows and the thickness of the rings, tree rings are used to learn about past climate and eruption. Researchers compared tree rings from around the world to determine that a volcano on Santorini probably erupted in 1560 B.C.

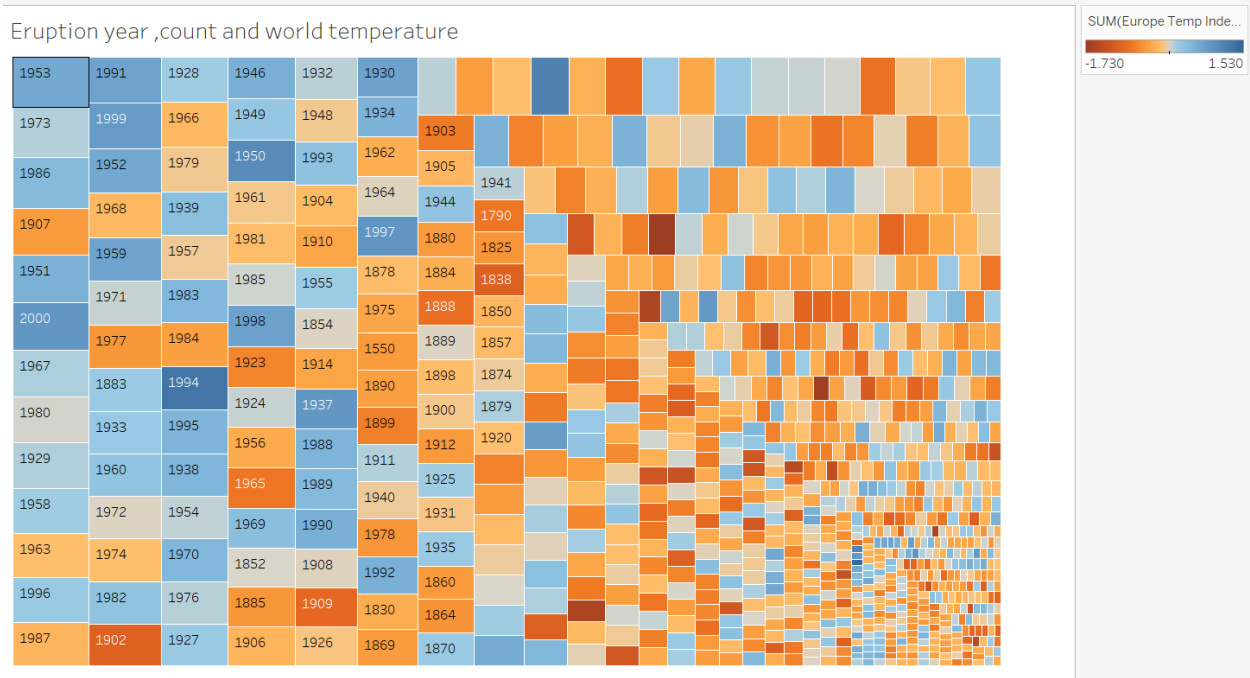
We need to merge two datasets (eruption and tree rings) to visualize the ring pattern over the years.



5. More frequent eruptions cause an increase in world temperatures.

Large-scale volcanic activity may last only a few days, but the massive outpouring of gases and ash can influence climate patterns for years which may increase the world temperature. Sulfuric gases convert to sulfate aerosols, sub-micron droplets containing about 75 percent sulfuric acid that stay in the stratosphere as long as 3 to 4 years; however, they also absorb heat radiated from the Earth, thereby warming the upper atmosphere (or stratosphere). major eruptions alter the Earth's radiative balance because volcanic aerosol clouds absorb terrestrial radiation, and scatter a significant amount of the incoming solar radiation, an effect known as "radiative forcing" that can last from two to three years following a volcanic eruption.

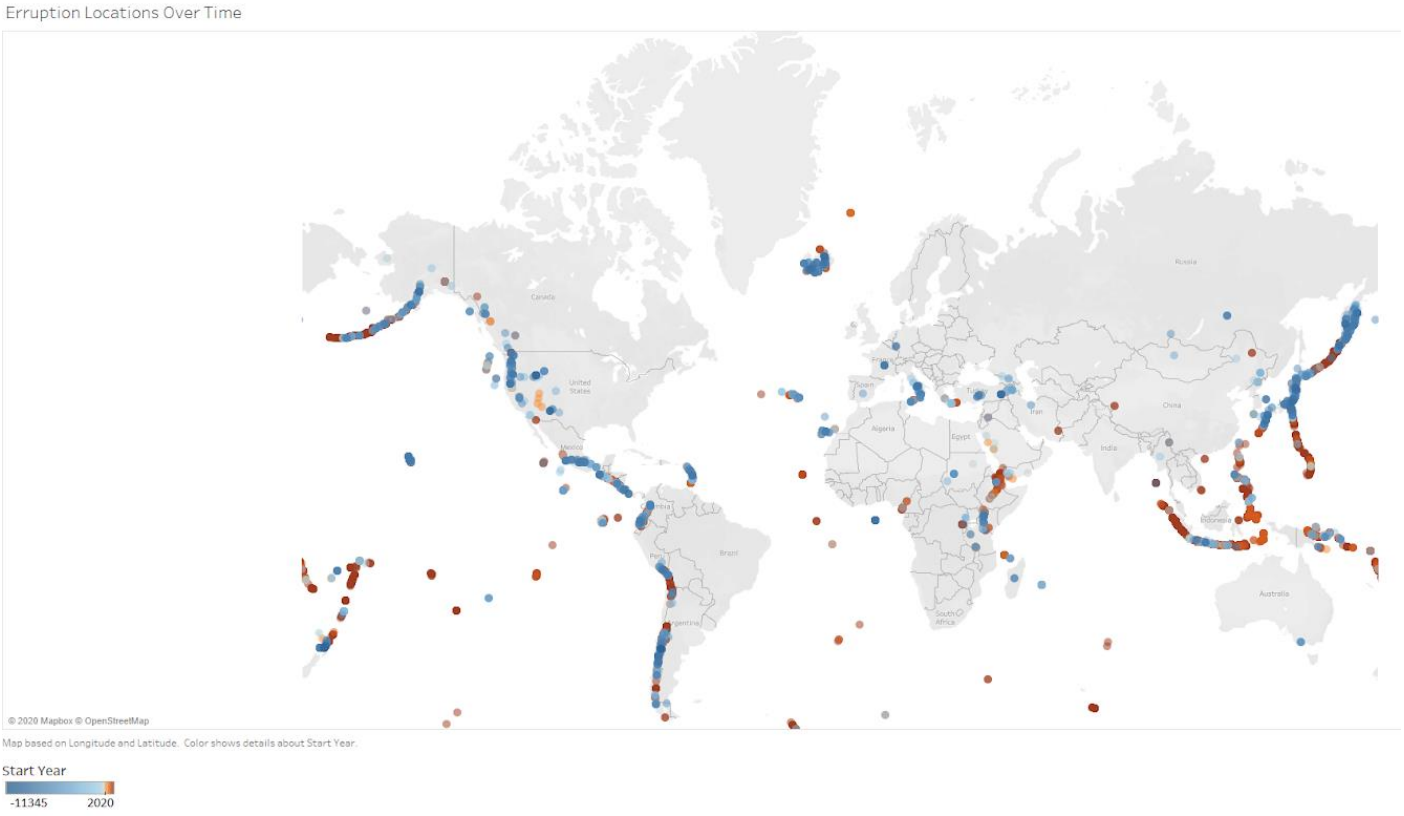
To visualize the world temperature, we need to plot between eruption year, volcanic number, and temperature, in addition, the sulfuric trend can also be shown by plotting another graph.



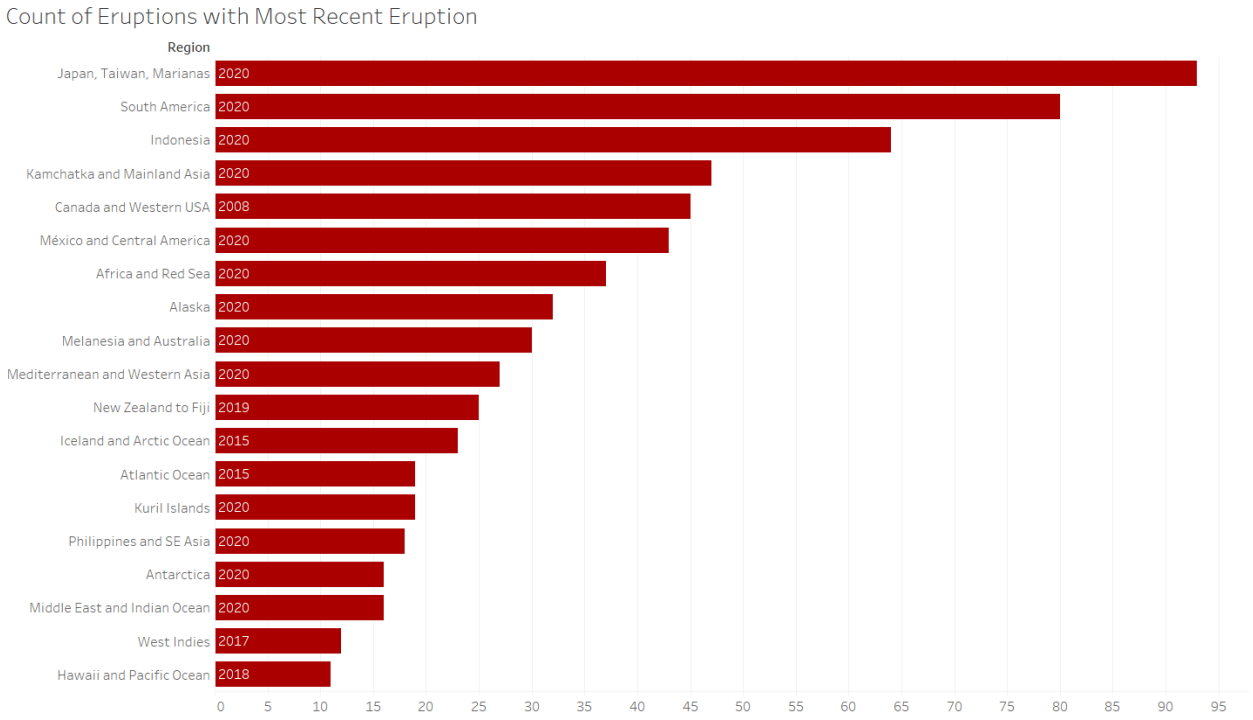
6. Areas of volcanic activity have changed over time.

This data contains records of volcanic activity dating back over 11,000 years BC to present day. As we have become more aware of volcanoes and their destructive behavior, our record keeping of their activity has increased. Our knowledge of the earliest volcanoes is mostly evidence based but recent eruptions have been observed and recorded. While we found that most volcanic eruptions occur along the intersections of tectonic plates, some boundaries are more active than others. The regions where a majority of volcanoes occur have seen an increase in activity in the past 1000 years. These regions include eastern parts of Asia and western areas of South America

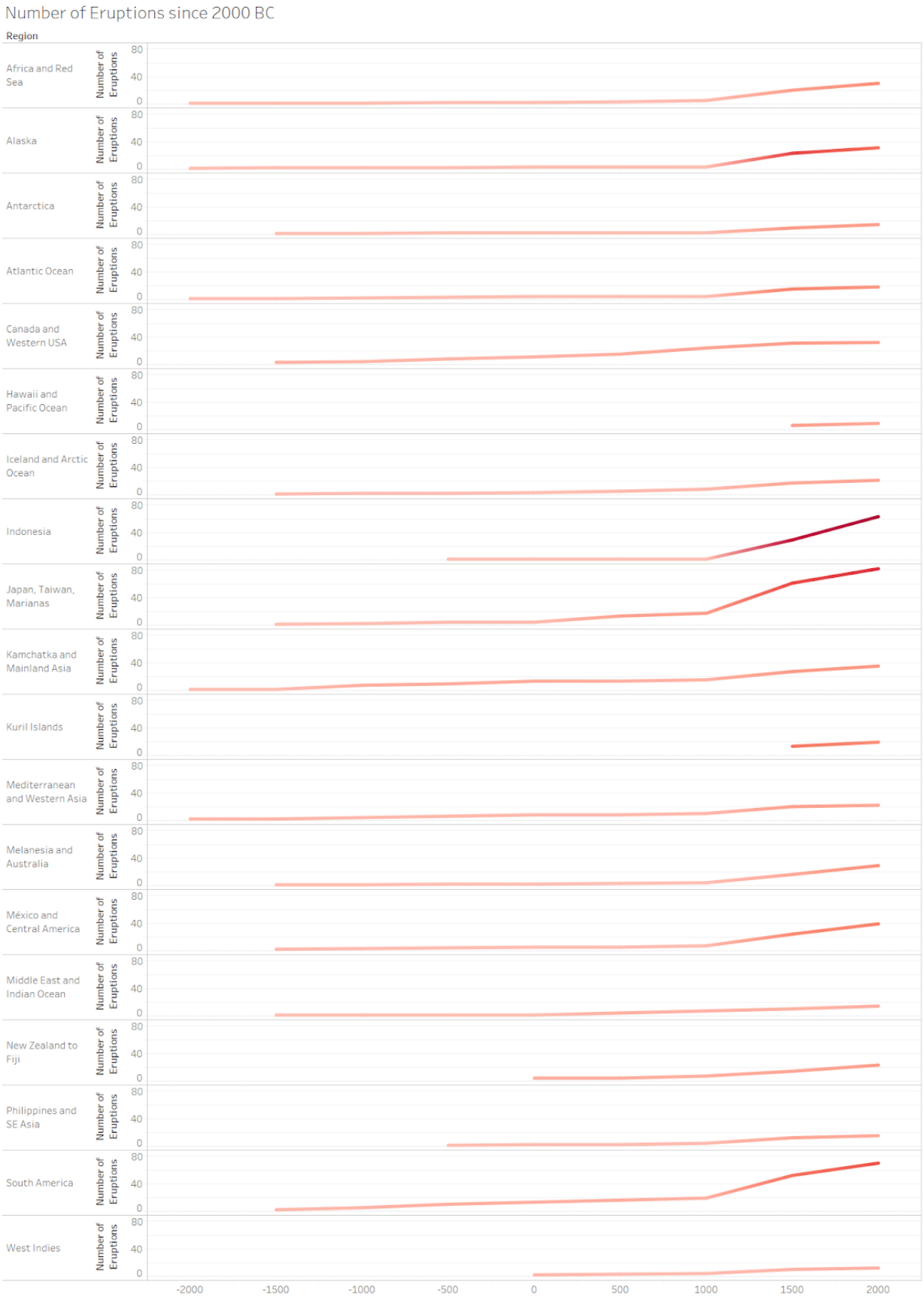
The graph below shows volcanic activity in different locations, with time encoded by the color of each point. Volcanoes that have been active in the last thousand years are represented by warmer colors. The more striking trend is that islands tend to have much more recent activity



than other areas. The Indian and Pacific ocean coastlines in the East have the most recent activity. This includes countries like Japan, Indonesia, and the Polynesian islands. There is also a lot of recent activity on the Aleut islands off of Alaska, as well as some activity in South America and The Middle East. Most of the eruptions occur along fault lines, but there little recorded activity along most underwater fault lines. This may simply be because activity is more difficult to measure there.



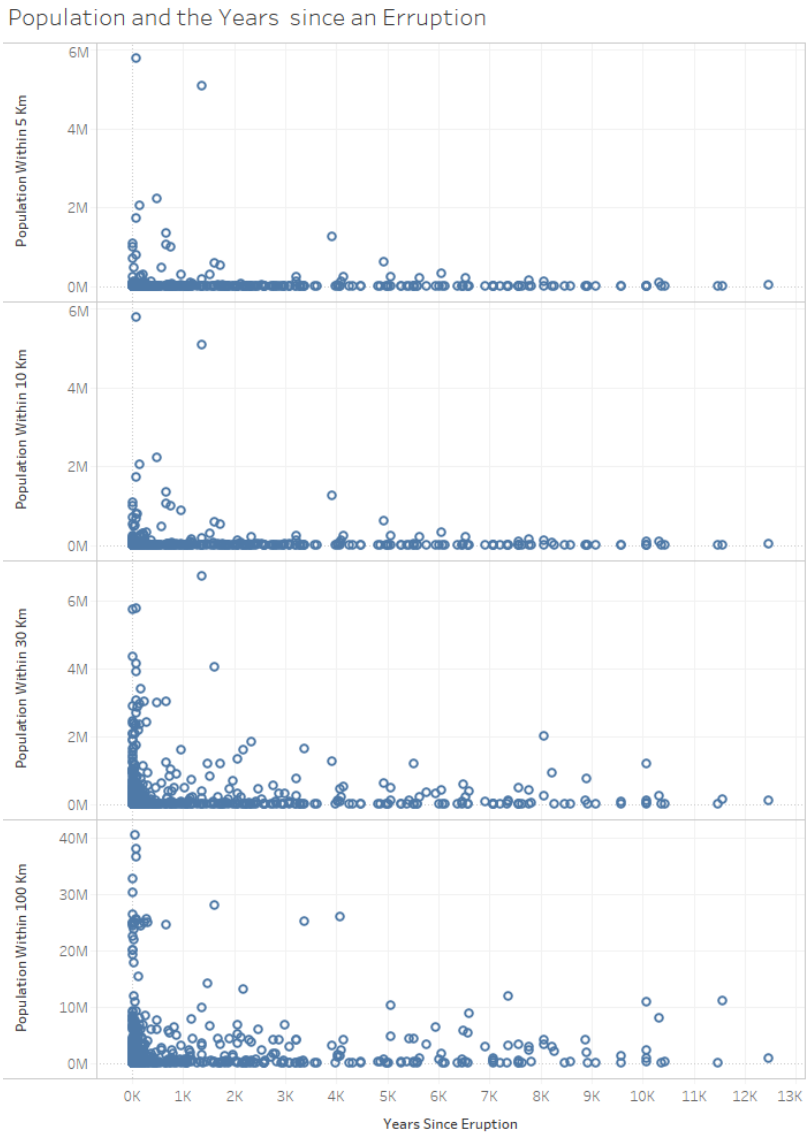
When looking at the recent eruption count by region, the absolute count is the highest in Japan, South America, Indonesia, Mainland Asia, Canada and the Western USA, as well as Africa and Alaska. Alaska, Indonesia, Japan, and South America have experienced the highest increases in eruptions since 2000 BC, as seen on the graph below.


Number of Eruptions since 2000 BC
Region
Africa and Red Sea
Alaska
Antarctica
Atlantic Ocean
Canada and Western USA
Hawaii and Pacific Ocean
Iceland and Arctic Ocean
Indonesia
Japan, Taiwan, Marianas
Kamchatka and Mainland Asia
Kuril Islands
Mediterranean and Western Asia
Melanesia and Australia
México and Central America
Middle East and Indian Ocean
New Zealand to Fiji
Philippines and SE Asia
South America
West Indies
Number of Eruptions
80
40
0
-2000
-1500
-1000
-500
0
500
1000
1500
2000

7. People tend to move away from volcanoes after increased activity

This graph shows the population around each volcano as it relates to the years since it last erupted.

If people were to move away from volcanoes following an eruption, we would expect the right side of the graph to have higher population numbers. However, the opposite appears to be true. More recently active volcanoes have higher population counts, which disproves our initial hypothesis.



Years Since Eruption vs. Population Within 5 Km, Population Within 10 Km, Population Within 30 Km and Population Within 100 Km.

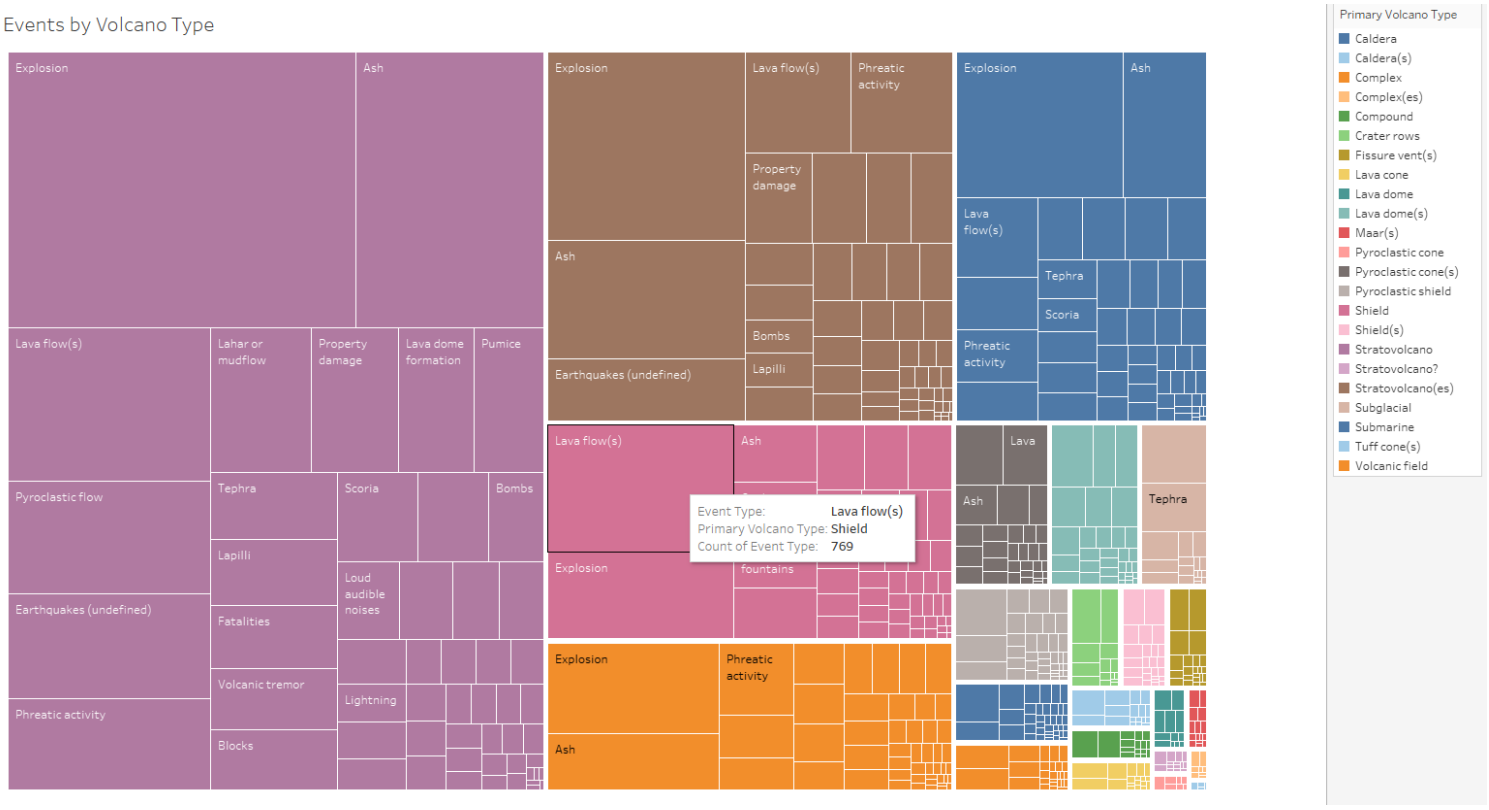
8. Certain events are better predictors of explosions than others.

Volcanologists often try and predict volcanic eruptions by analyzing different predictors. Volcanoes that are about to erupt show several indications, such as increases in the frequency and intensity of earthquakes and tremors, steaming, volcanic gas, hot and enlarged areas on the ground, swelling of the surface, etc. Since the events dataset only contains events related to eruptions that occurred (as opposed to activity that did not lead to an eruption), we could not build a model to predict what the major factors or eruptions were. However, we summarized what type of events happen prior to eruptions and where they are occurring.

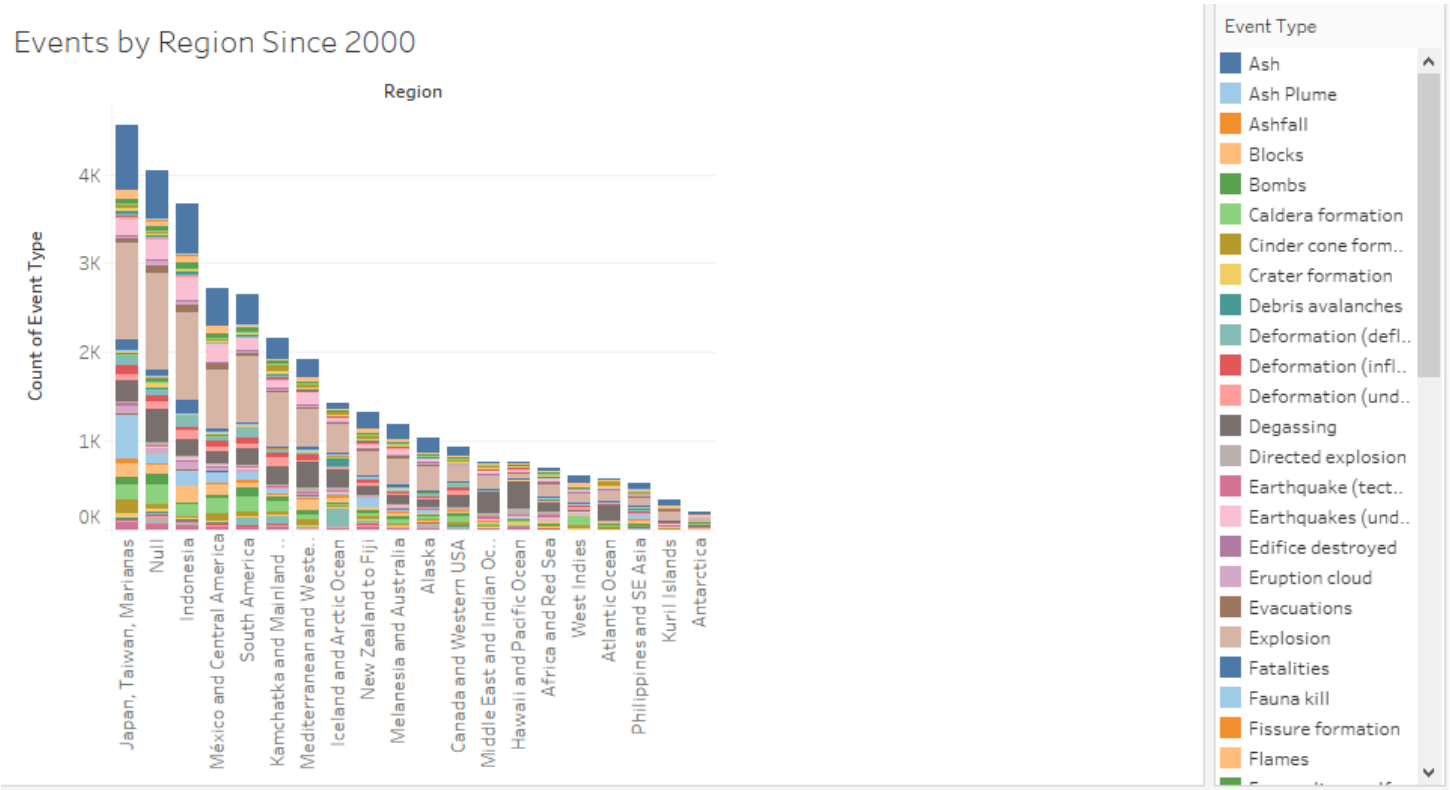
In the below geographical visualization, we can see the aggregated event types per location and filter based on these event types. For instance, if we wanted to know which volcanoes are producing lava lakes upon eruption, we can filter on the lava lake event type to see Kilauea has produced the most amount of lava lake events, while there are clusters of other volcanoes which have this event type too, such as the grouping around Iceland. Other types of

events, such as loud noises, often happened in Indonesia. Producing this visualization allows the user to explore the may event types and look around the map to see what they are interested in.

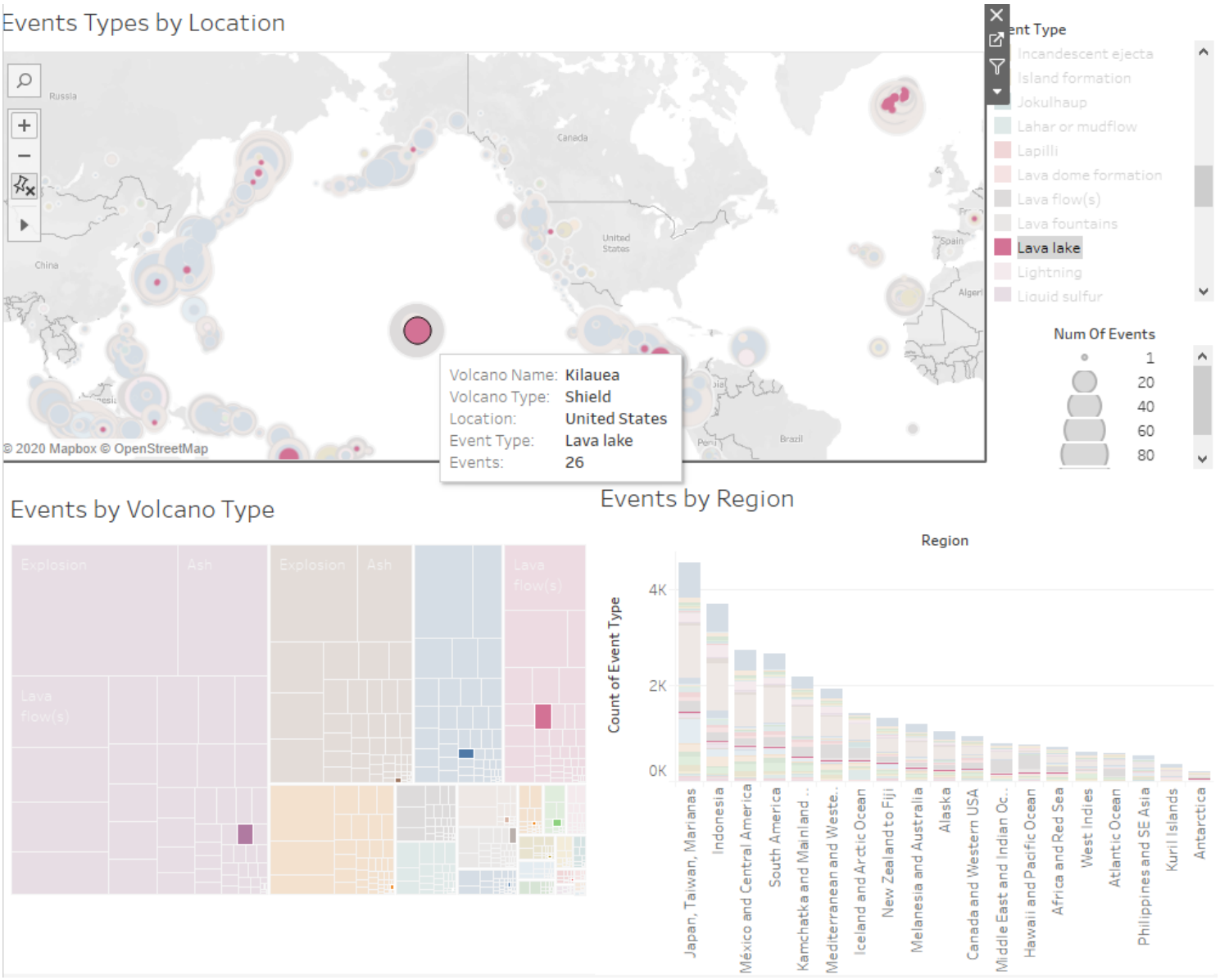
For a breakdown of events by volcano type, we produced a treemap and grouped by the volcano type. From this, you can see Stratovolcanoes (left group in purple) are the most common volcanoes. In addition, the most common event type for each volcano type was often Explosion, Ash, and Lava flows. Some volcano types, such as Shield, have Lava Flows as their most common event type.



Next, we created a stacked bar chart to show the breakdown of events by region. This allows you to hover over each region and see with color and labels from the tooltip, how frequent that event type is for the region you are looking at.

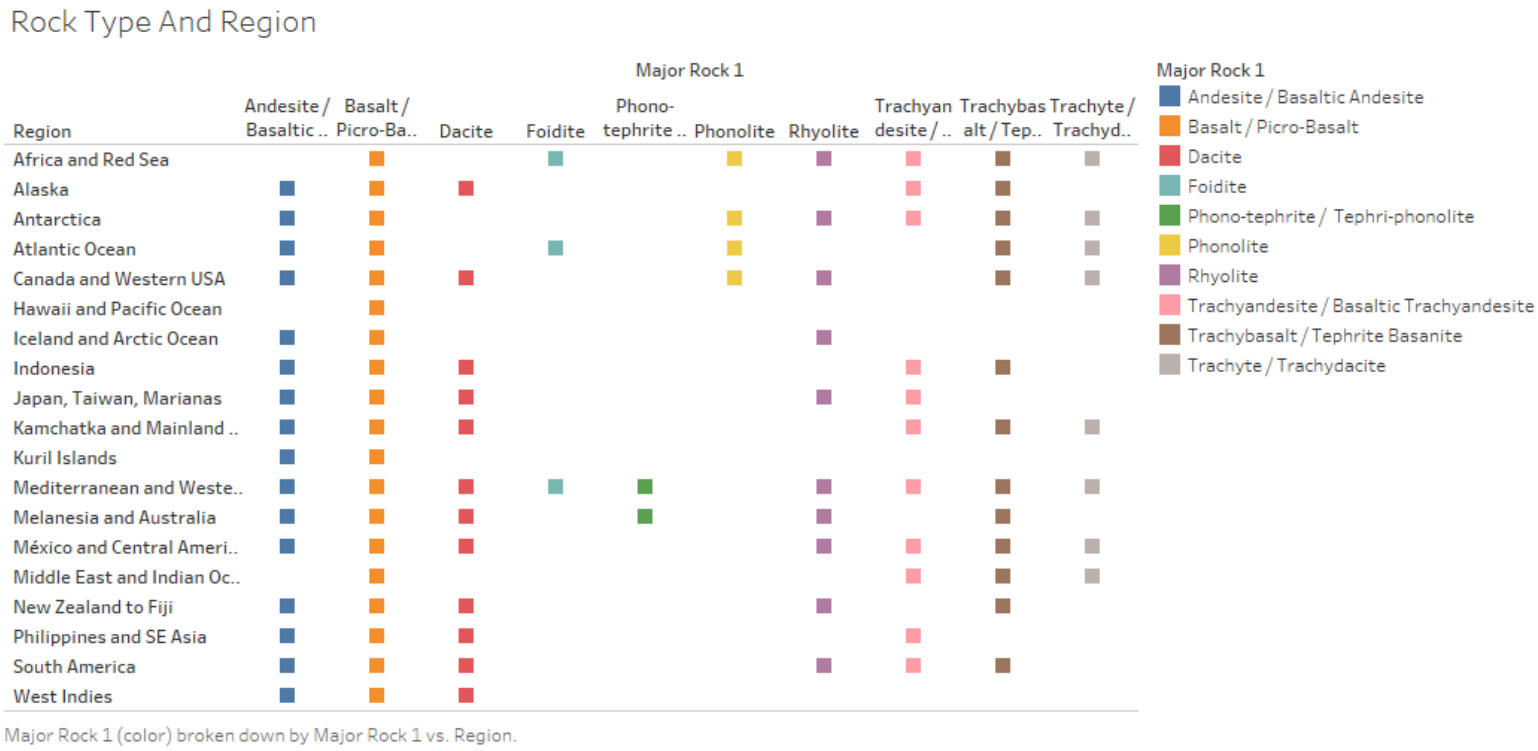


To combine everything into one, we created a dashboard with the above 3 visualizations. Using this dashboard, you can filter on the event type. In the example below, we are filtering again on Lava Lakes. From this we can see geographically where this event type is often found on the map. Then we can look at the treemap below and see for each volcano type, the frequency with respect to other event types. Lastly, you can then look at the bottom right to get an idea of the regional occurrence of the event you are looking at.


Events Types by Location
Russia
Canada
United States
China
Spain
Brazil
© 2020 Mapbox © OpenStreetMap
Volcano Name: Kilauea
Volcano Type: Shield
Location: United States
Event Type: Lava lake
Events: 26
ent Type
Incandescent ejecta
Island formation
Jokulhaup
Lahar or mudflow
Lapilli
Lava dome formation
Lava flow(s)
Lava fountains
Lava lake
Lightning
Liquid sulfur
Num Of Events
1
20
40
60
80
Events by Volcano Type
Explosion
Ash
Lava flow(s)
Events by Region
Region
Count of Event Type
4K
2K
0K
Japan, Taiwan, Marianas
Indonesia
México and Central America
South America
Kamchatka and Mainland ..
Mediterranean and Weste..
Iceland and Arctic Ocean
New Zealand to Fiji
Melanesia and Australia
Alaska
Canada and Western USA
Middle East and Indian Oc..
Hawaii and Pacific Ocean
Africa and Red Sea
West Indies
Atlantic Ocean
Philippines and SE Asia
Kuril Islands
Antarctica

9. Rock type varies by location.

The volcano data contained information on the primary type of rock present, which we believed would vary depending on the location of the volcano. Andesite and Basalt are present in every region, while Foldite and Phono-tephrite are only found in a handful of locations.



Conclusion

After analyzing the data, we found that most volcanoes occur along fault lines, and that islands are more likely to exhibit recent volcanic activity. Since many faults occur along the coastlines of continents, where population density is high and there is access to ports, people do not tend to move following volcanic activity in the way we expected. Japan, South America, and Indonesia have the highest occurrence of recent activity. While nearby population counts were affected, we did find that volcanic activity is correlated with an increase in global temperatures.

Predicting volcanic eruptions with this dataset was not possible, but we found the most common events occurring with eruptions are explosions, ash, lava flows, and earthquakes. In addition, events are often linked to geographic positions and volcano types. Observing all this information helps us gain a better understanding of trends in volcanic activity.