

Data Exploration with SQL

In this exercise, we will be exploring the Volcanic Eruptions dataset that has 822 records for different volcanic eruptions around the world. I used MySQL to load the dataset and execute the following queries.

For ease of use, I renamed the table from **volcanic_eruptions** to simply **volcano**. Below are ten records displayed from the **volcano** table after importing the data from the CSV file.

	recno	Year	Month	Day	Name	Location	Country	Latitude	Longitude	Elevation	Type	VEI	Deaths
►	1	-1750			Veniaminof	Alaska Peninsula	United States	56.17	-159.38	2507	Stratovolcano	6	
	2	-1645			Aniakchak	Alaska Peninsula	United States	56.88	-158.17	1341	Caldera	6	
	3	-1610			Santorini	Greece	Greece	36.404	25.396	329	Shield volcano	7	
	4	-1550			Redoubt	Alaska-SW	United States	60.48	-152.75	3108	Stratovolcano		
	5	-1460			Taupo	New Zealand	New Zealand	-38.82	176	760	Caldera	6	
	6	-1370			Pago	New Britain-SW Pac	Papua New Guinea	-5.58	150.52	742	Caldera	6	
	7	-1050			Pinatubo	Luzon-Philippines	Philippines	15.13	120.35	1486	Stratovolcano	6	
	8	-1010			Tungurahua	Ecuador	Ecuador	-1.467	-78.442	5023	Stratovolcano	5	
	9	-455			Cotopaxi	Ecuador	Ecuador	-0.677	-78.436	5911	Stratovolcano	4	
	10	-250			Raoul Island	Kermadec Is	New Zealand	-29.27	-177.92	516	Stratovolcano	6	

Query the database and *interpret* the results, displaying:

- Number of Volcanoes per Country
 - By examining the query results below, we find that Indonesia and Japan are the Top 2 countries with the highest number of volcanoes. It can also be noted that most countries on this list only have one volcano.

Country	Volcano_Count	Country	Volcano_Count
► Indonesia	44	Cameroon	2
Japan	37	Congo, DRC	2
United States	22	Tanzania	2
Papua New Guinea	17	Eritrea	2
Russia	14	Costa Rica	2
Iceland	11	Canada	1
Philippines	10	Cape Verde	1
Chile	9	China	1
Ecuador	7	Comoros	1
Nicaragua	6	Martinique	1
Colombia	5	Guadeloupe	1
New Zealand	5	Reunion	1
Italy	5	Greece	1
Mexico	5	Samoa	1
Vanuatu	5	Saudi Arabia	1
El Salvador	4	Netherlands	1
Ethiopia	4	Montserrat	1
Portugal	4	St. Kitts & Nevis	1
Tonga	3	St. Vincent & the Grenadines	1
Pacific Ocean	3	Taiwan	1
Peru	3	Trinidad	1
Solomon Is.	3	Turkey	1
Spain	3	Antarctica	1
Yemen	3	Vietnam	1
Guatemala	3	North Korea	1

- Number of Volcanoes of Each Type

- This table tells us that the Stratovolcano is the most common volcano type according to this dataset, with 169 unique volcanos. Eight other volcano types are tied with only one volcano per type.

	Type	Volcano_Count
▶	Stratovolcano	169
	Complex volcano	23
	Caldera	21
	Shield volcano	18
	Submarine volcano	6
	Pyroclastic shield	4
	Lava dome	3
	Pyroclastic cone	3
	Cinder cone	2
	Compound volcano	2
	Fissure vent	2
	Volcanic field	2
	Crater rows	1
	Lava cone	1
	Maar	1
	Mud volcano	1
	Pumice cone	1
	Subglacial volcano	1
	Submarine volcanoes	1
	Tuff cone	1

- VEI Summary Statistics

- The **VEI** summary statistics were a bit different and difficult to calculate than other columns. We notice a min of 0, an average of 2.2, a median of 2, and a maximum of 7 VEI. This tells us that the maximum of 7 VEI is likely an uncommon occurrence for a volcanic eruption.
- However, one critical flaw is that this column had 0 and NULL values in it, so the data was not complete and accurate.

	min_vei	avg_vei	median_vei	max_vei
▶	0	2.22	2	7

- Elevation summary statistics

- Summary statistics for the **Elevation** column reveal a min of -642, an average of 1997, a median of 1725, and a maximum of 5967. This shows us that most eruptions occur 1000-2000 feet above sea level and even underwater, but that it is rare for a very high volcano (>5000 ft) to erupt.

	min_elevation	avg_elevation	median_elevation	max_elevation
▶	-642	1997	1725	5967

- Most Active Volcanoes (Repeated Eruptions)

- Lastly, we queried the dataset to find the most active volcanoes – or those with repeated eruptions. To do this, we pulled the Volcano Name and the Eruption Count. Results reveal that the Merapi volcano is the most active volcano, with 27 total eruptions according to this dataset. Next is Etna followed by Semeru. In the end, there are five volcanoes considered the least active with only 4 eruptions each.

Volcano	Eruption_Count	Volcano	Eruption_Count
Merapi	27	Grimsvoth	7
Etna	25	Paluweh	7
Semeru	23	Myojun Knoll	6
Mayon	19	Banda Api	6
Kelut	18	Taal	6
Asama	18	Bardarbunga	6
Vesuvius	17	Soufriere Hills	6
Aso	16	Awu	6
Hekla	16	Karthala	6
Katla	15	Rabaul	6
Kilauea	15	Cotopaxi	6
Sakura-jima	14	Unzen	5
Tungurahua	11	Camiguin	5
Fuego	11	Aoba	5
Stromboli	11	Hakone	5
Karangetang [Api Siau]	10	Calbuco	5
Gamalama	10	Yasur	5
Sinabung	10	Raung	5
Dieng Volc Complex	10	Iliwerung	5
Kirishima	9	Pinatubo	5
Manam	9	Okataina	5
Santorini	8	Tinakula	4
Krakatau	8	Chichon, El	4
Usu	7	Bulusan	4
Grimsvoth	7	Taupo	4
		Colima	4