VIETNAM NATIONAL UNIVERSITY HCMC
UNIVERSITY OF INFORMATION TECHNOLOGY

Subject: Decision support and business intelligence applications

Topic: Analysis the Global Historical Tsunami Database

Lecturer: Do Phuc
Lecturer: Nguyen Thi Kim Phung

Group: 2

Members:
- Pham Thuy Dung - 20521214
- Dau Dinh Quang Anh - 20521059

COMMENT

Table of Contents

1 Topic Overview

This report conducts research on the causes, consequences, and damage of tsunamis based on the Global Historical Tsunami Database.

This dataset represents historical tsunamis and related information, providing data on over 2,400 tsunamis from 2100 BC to the present in the Atlantic, Indian, and Pacific Oceans, as well as the Mediterranean and Caribbean Seas. This dataset download from NCEI Global Historical Hazard Database (noaa.gov)

2 Dictionary of Data

Attribute	Description
Year	The year in which the event occurred.
Mo (Month)	The month in which the event occurred.
Dy (Day)	The day on which the event occurred.
Hr (Hour)	The hour at which the event occurred.
Mn (Minute)	The minute at which the event occurred.
Sec (Second)	The second at which the event occurred.
Tsunami Event Validity	Valid values: -1 to 4 Validity of the actual tsunami occurrence is indicated by a numerical rating of the reports of that event: -1: erroneous entry 0: event that only caused a seiche or disturbance in an inland river/lake 1: very doubtful tsunami 2: questionable tsunami 3: probable tsunami 4: definite tsunami
Tsunami Cause Code	Valid values: 0 to 11 The source of the tsunami: 0: Unknown 1: Earthquake 2: Questionable Earthquake 3: Earthquake and Landslide 4: Volcano and Earthquake 5: Volcano, Earthquake, and Landslide 6: Volcano 7: Volcano and Landslide 8: Landslide 9: Meteorological 10: Explosion 11: Astronomical Tide
Earthquake Magnitude	Valid values: 0.0 to 9.9

	The value in this column contains the primary earthquake magnitude. Magnitude measures the energy released at the source of the earthquake. Magnitude is determined from measurements on seismographs. For pre-instrumental events, the magnitudes are derived from intensities. There are several different scales for measuring earthquake magnitudes. The primary magnitude is chosen from the available magnitude scales in this order: • Mw Magnitude • Ms Magnitude • Mb Magnitude • Ml Magnitude • Mfa Magnitude • Unknown Magnitude
Vol	Volcano identification number if a volcanic eruption caused the tsunami.
More Info	Additional information or notes about the event.
Deposits	Information about deposits left by the tsunami.
Country	The Country where the tsunami source occurred (For example enter: Japan or Russia).
Location Name	The Country, State, Province or Island where the tsunami source occurred (For example enter: Japan or Honshu
Latitude (kinh độ)	Valid values: **-90 to 90** Latitude: 0 to 90 (Northern Hemisphere) -90 to 0 (Southern Hemisphere)
Longitude (vĩ độ)	Valid values: **-180 to 180** Longitude: 0 to 180 (Eastern Hemisphere) -180 to 0 (Western Hemisphere)
Maximum Water Height (m)	Maximum Water Height (m) • The maximum water height above sea level in meters for this event. If the type of measurement of the runup was a: • Tide Gauge - half of the maximum height (minus the normal tide) of a tsunami wave recorded at the coast by a tide gauge.

	• Runup Height - the maximum elevation the wave reaches at the maximum inundation.
Number of Runups	The total number of runups link will display the runup locations associated with a particular tsunami event.
Tsunami Magnitude (Abe)	Tsunami magnitude scale by Abe.
Tsunami Magnitude (Iida)	Tsunami magnitude scale by Iida.
Tsunami Intensity	Intensity of the tsunami.
Deaths	Whenever possible, numbers of deaths are listed.
Death Description	Valid values: 0 to 4 When a description was found in the historical literature instead of an actual number of deaths, this value was coded and listed in the Deaths column. If the actual number of deaths was listed, a descriptor was also added for search purposes. 0: None 1: Few (~1 to 50 deaths) 2: Some (~51 to 100 deaths) 3: Many (~101 to 1000 deaths) 4: Very many (over 1000 deaths)
Missing	Number of people missing after the tsunami.
Missing Description	Description of missing people.
Injuries	Whenever possible, numbers of injuries from the tsunami are listed.
Injuries Description	Valid values: 0 to 4 When a description was found in the historical literature instead of an actual number of injuries, this value was coded and listed in the Injuries column. If the actual number of injuries was listed, a descriptor was also added for search purposes. 0: None 1: Few (~1 to 50 injuries) 2: Some(~51 to 100 injuries) 3: Many (~101 to 1000 injuries) 4: Very many (over 1000 injuries)

Damage ($Mil)	The value in the Damage column should be multiplied by 1,000,000 to obtain the actual dollar amount.
Damage Description	Valid values: 0 to 4 For those events not offering a monetary evaluation of damage, the following five-level scale was used to classify damage (1990 dollars) and was listed in the Damage column. If the actual dollar amount of damage was listed, a descriptor was also added for search purposes. 0: NONE 1: LIMITED (roughly corresponding to less than $1 million) 2: MODERATE (~$1 to $5 million) 3: SEVERE (~$5 to $25 million) 4: EXTREME (~$25 million or more)
Houses Destroyed	Whenever possible, numbers of houses destroyed are listed.
Houses Destroyed Description	Valid values: 0 to 4 For those events not offering an exact number of houses damaged, the following four-level scale was used to classify the damage and was listed in the Houses Destroyed column. If the actual number of houses destroyed was listed, a descriptor was also added for search purposes. 0: None 1: Few (~1 to 50 houses) 2: Some (~51 to 100 houses) 3: Many (101 to 1000 houses) 4: Very Many (~over 1000 houses)
Houses Damaged	Number of houses damaged.
Houses Damaged Description	Description of the damage to houses.
Total Deaths	Whenever possible, total number of deaths from the tsunami and secondary effects are listed.
Total Death Description	Valid values: 0 to 4 When a description was found in the historical literature instead of an actual number of deaths, this value was coded and listed in the Deaths column. If the actual number of deaths was listed, a descriptor was also added for search purposes.

	0: None 1: Few (~1 to 50 deaths) 2: Some (~51 to 100 deaths) 3: Many (~101 to 1000 deaths) 4: Very many (over 1000 deaths)
Total Missing	Total number of missing including secondary effects.
Total Missing Description	Description of total missing.
Total Injuries	Total number of injuries including secondary effects.
Total Injuries Description	Description of total injuries.
Total Damage ($Mil)	Total Damage in Millions of Dollars from the Tsunami and the Source Event
Total Damage Description	Valid values: 0 to 4 For those events not offering a monetary evaluation of damage, the following five-level scale was used to classify damage (1990 dollars) and was listed in the Damage column. If the actual dollar amount of damage was listed, a descriptor was also added for search purposes. 0: NONE 1: LIMITED (roughly corresponding to less than $1 million) 2: MODERATE (~$1 to $5 million) 3: SEVERE (~$5 to $25 million) 4: EXTREME (~$25 million or more)
Total Houses Destroyed	Total number of houses destroyed.
Total Houses Destroyed Description	Description of total houses destroyed.
Total Houses Damaged	Total number of houses damaged.
Total Houses Damaged Description	Description of total houses damaged.

3 Data Warehouse Design

DimTime
- [ID Time]
- Hr
- Mn
- Sec

DimDate
- [ID Date]
- Year
- Mo
- Dy

DimLocation
- [ID Location]
- Country
- [Location Name]
- Latitude
- Longitude
- [Maximum Water Height (m)]
- [Number of Runups]
- [Tsunami Magnitude (Abe)]
- [Tsunami Magnitude (Iida)]
- [Tsunami Intensity]

Fact
- [ID Fact]
- [ID Date]
- [ID Time]
- [ID Description]
- [ID Location]
- [ID Detail Damage]
- [ID Total Damage]

DimTotalDamage
- [ID Total Damage]
- [Total Deaths]
- [Total Death Description]
- [Total Missing]
- [Total Missing Description]
- [Total Injuries]
- [Total Injuries Description]
- [Total Damage ($Mil)]
- [Total Damage Description]
- [Total Houses Destroyed]
- [Total Houses Destroyed Description]
- [Total Houses Damaged]
- [Total Houses Damaged Description]

DimDetailDamage
- [ID Detail Damage]
- Deaths
- [Death Description]
- Missing
- [Missing Description]
- Injuries
- [Injuries Description]
- [Damage ($Mil)]
- [Damage Description]
- [Houses Destroyed]
- [Houses Destroyed Description]
- [Houses Damaged]
- [Houses Damaged Description]

DimDescription
- [ID Description]
- [Tsunami Event Validity]
- [Tsunami Cause Code]
- [Earthquake Magnitude]
- Vol
- [More Info]
- Deposits

4 Buid Data Warehouse (SSIS)
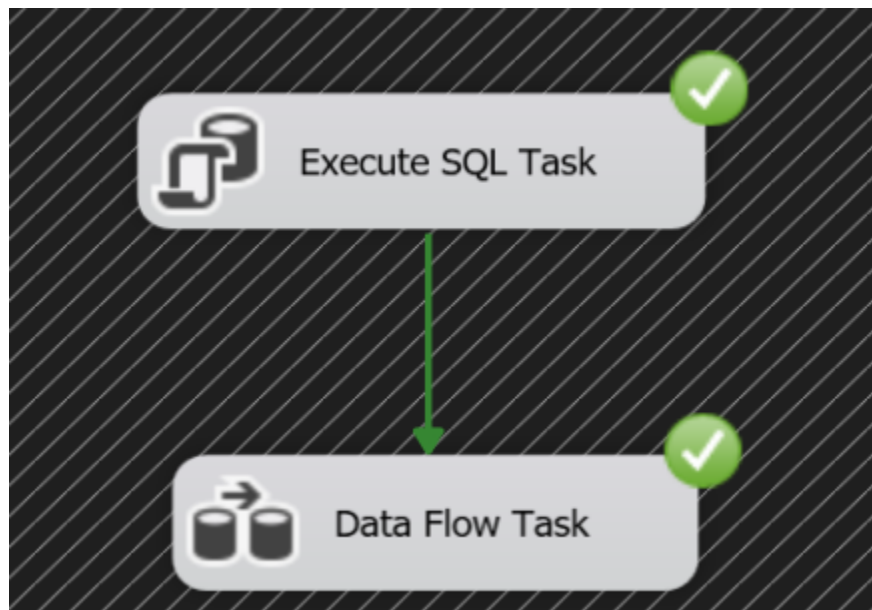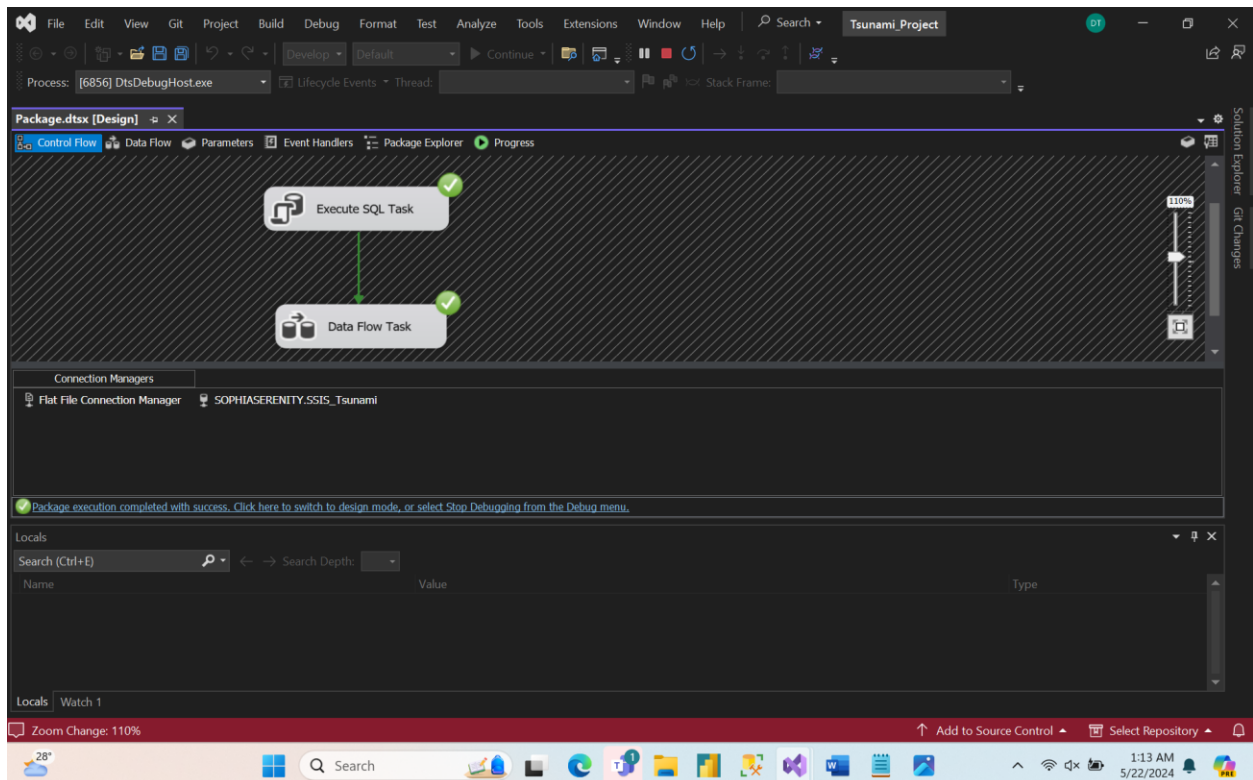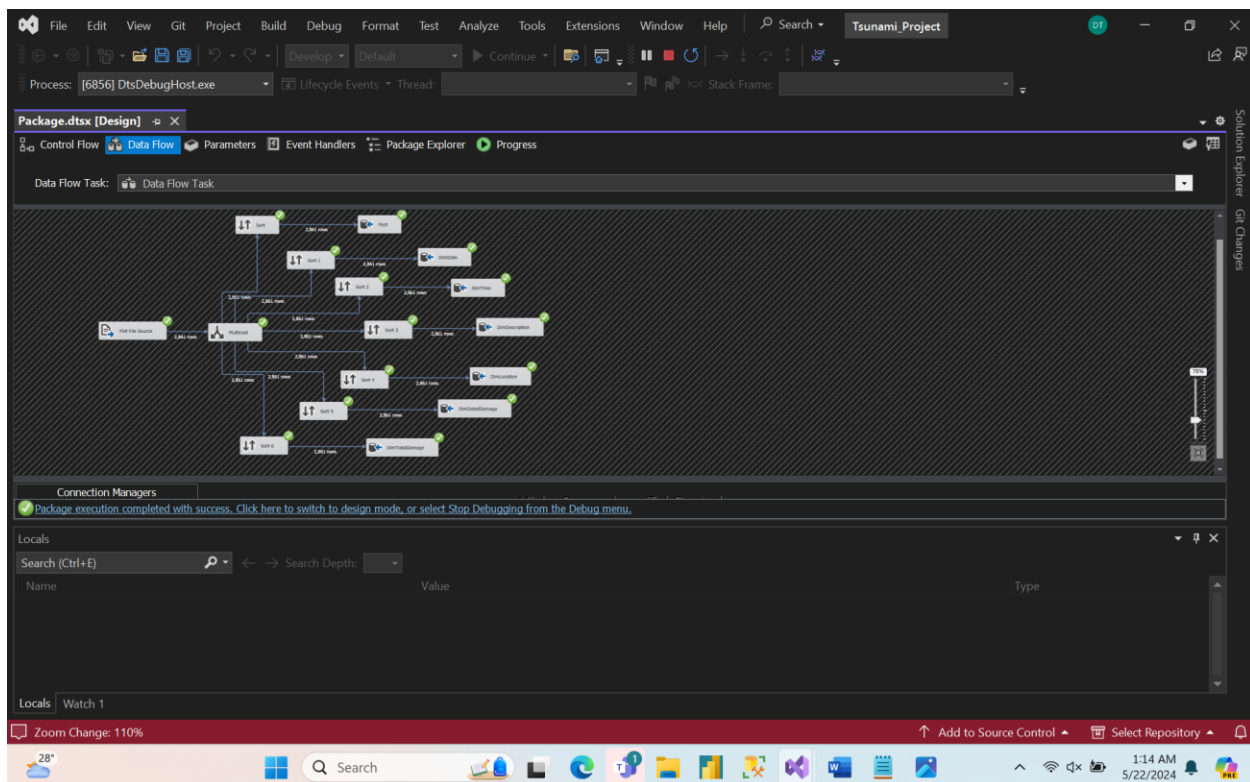
The source data used for analysis is a csv file, so we use flat file sources to integrate the data, use sort to remove and order the lines.

5 Data Warehouse Analysis (SSAS)

Data Source View

DimLocation
- 🔑 ID Location
- Country
- Location Name
- Latitude
- Longitude
- Maximum Water Height (m)
- Number of Runups
- Tsunami Magnitude (Abe)
- Tsunami Magnitude (Iida)
- Tsunami Intensity

DimDescription
- 🔑 ID Description
- Tsunami Event Validity
- Tsunami Cause Code
- Earthquake Magnitude
- Vol
- More Info
- Deposits

DimDate
- 🔑 ID Date
- Year
- Mo
- Dy

Fact
- 🔑 ID Fact
- ID Date
- ID Time
- ID Description
- ID Location
- ID Detail Damage
- ID Total Damage

DimTime
- 🔑 ID Time
- Hr
- Mn
- Sec

DimDetailDamage
- 🔑 ID Detail Damage
- Deaths
- Death Description
- Missing
- Missing Description
- Injuries
- Injuries Description
- Damage ($Mil)
- Damage Description
- Houses Destroyed
- Houses Destroyed Description
- Houses Damaged
- Houses Damaged Description

DimTotalDamage
- 🔑 ID Total Damage
- Total Deaths
- Total Death Description
- Total Missing
- Total Missing Description
- Total Injuries
- Total Injuries Description
- Total Damage ($Mil)
- Total Damage Description
- Total Houses Destroyed
- Total Houses Destroyed Description
- Total Houses Damaged
- Total Houses Damaged Description

5.1 Create Data Warehouse

5.2 Create Data Source View

DimDescription
- ID Description
- Tsunami Event Validity
- Tsunami Cause Code
- Earthquake Magnitude
- Vol
- More Info
- Deposits

DimTotalDamage
- ID Total Damage
- Total Deaths
- Total Death Description
- Total Missing
- Total Missing Description
- Total Injuries
- Total Injuries Description
- Total Damage ($Mil)
- Total Damage Description
- Total Houses Destroyed
- Total Houses Destroyed Description
- Total Houses Damaged
- Total Houses Damaged Description

DimLocation
- ID Location
- Country
- Location Name
- Latitude
- Longitude
- Maximum Water Height (m)
- Number of Runups
- Tsunami Magnitude (Abe)
- Tsunami Magnitude (Iida)
- Tsunami Intensity

Fact
- ID Fact
- ID Date
- ID Time
- ID Description
- ID Location
- ID Detail Damage
- ID Total Damage

DimTime
- ID Time
- Hr
- Mn
- Sec

DimDetailDamage
- ID Detail Damage
- Deaths
- Death Description
- Missing
- Missing Description
- Injuries
- Injuries Description
- Damage ($Mil)
- Damage Description
- Houses Destroyed
- Houses Destroyed Description
- Houses Damaged
- Houses Damaged Description

DimDate
- ID Date
- Year
- Mo
- Dy

5.3 Create Data Cube

5.4 Edit the Dimension

5.4.1 DimDate

5.4.2 DimTime

Attributes	Hierarchies	Data Source View

Attributes

- Dim Time
  - Hr
  - ID Time
  - Mn
  - Sec

Hierarchies

To create a new hierarchy, drag an attribute here.

Data Source View
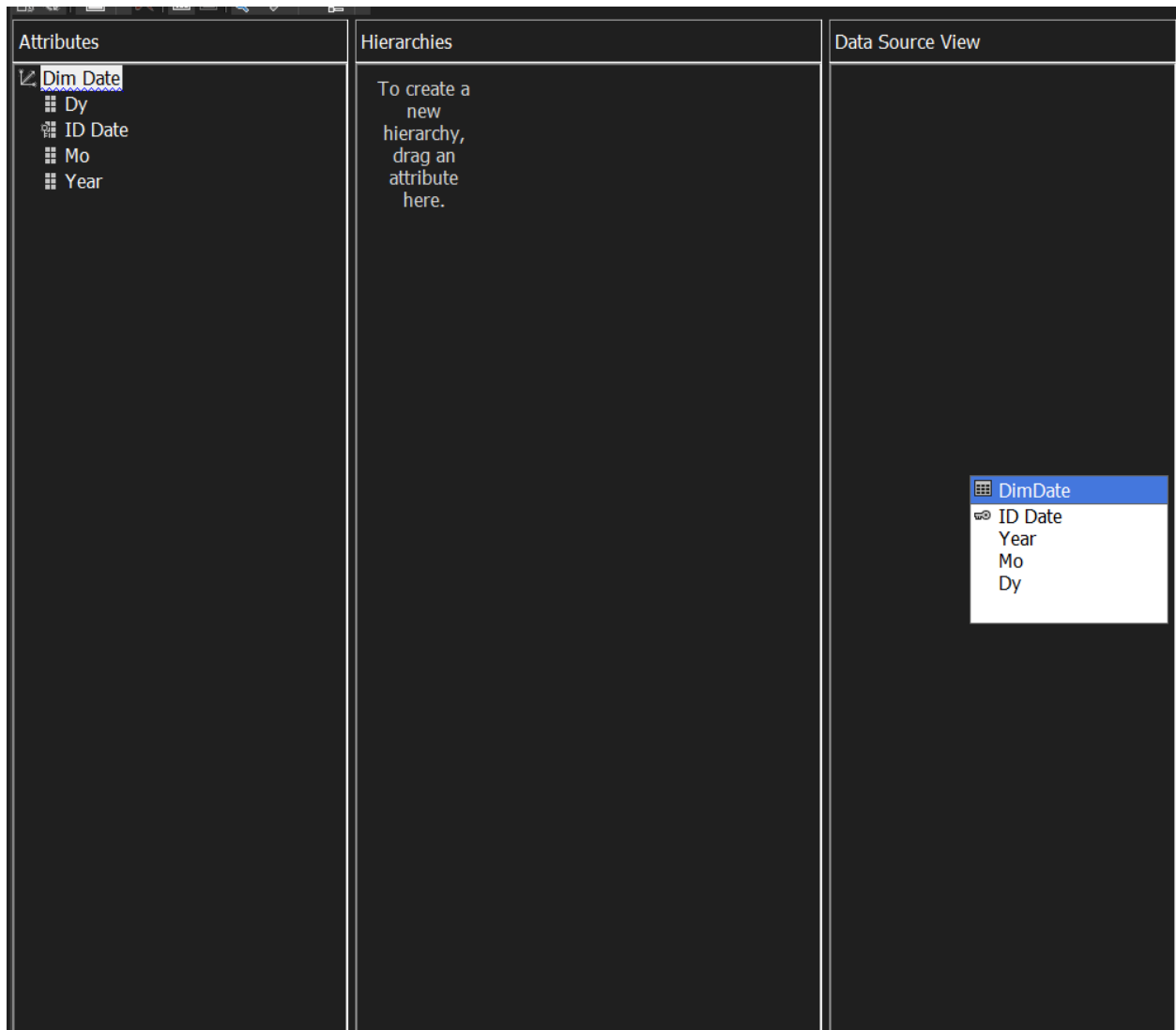
- DimTime
  - ID Time
  - Hr
  - Mn
  - Sec
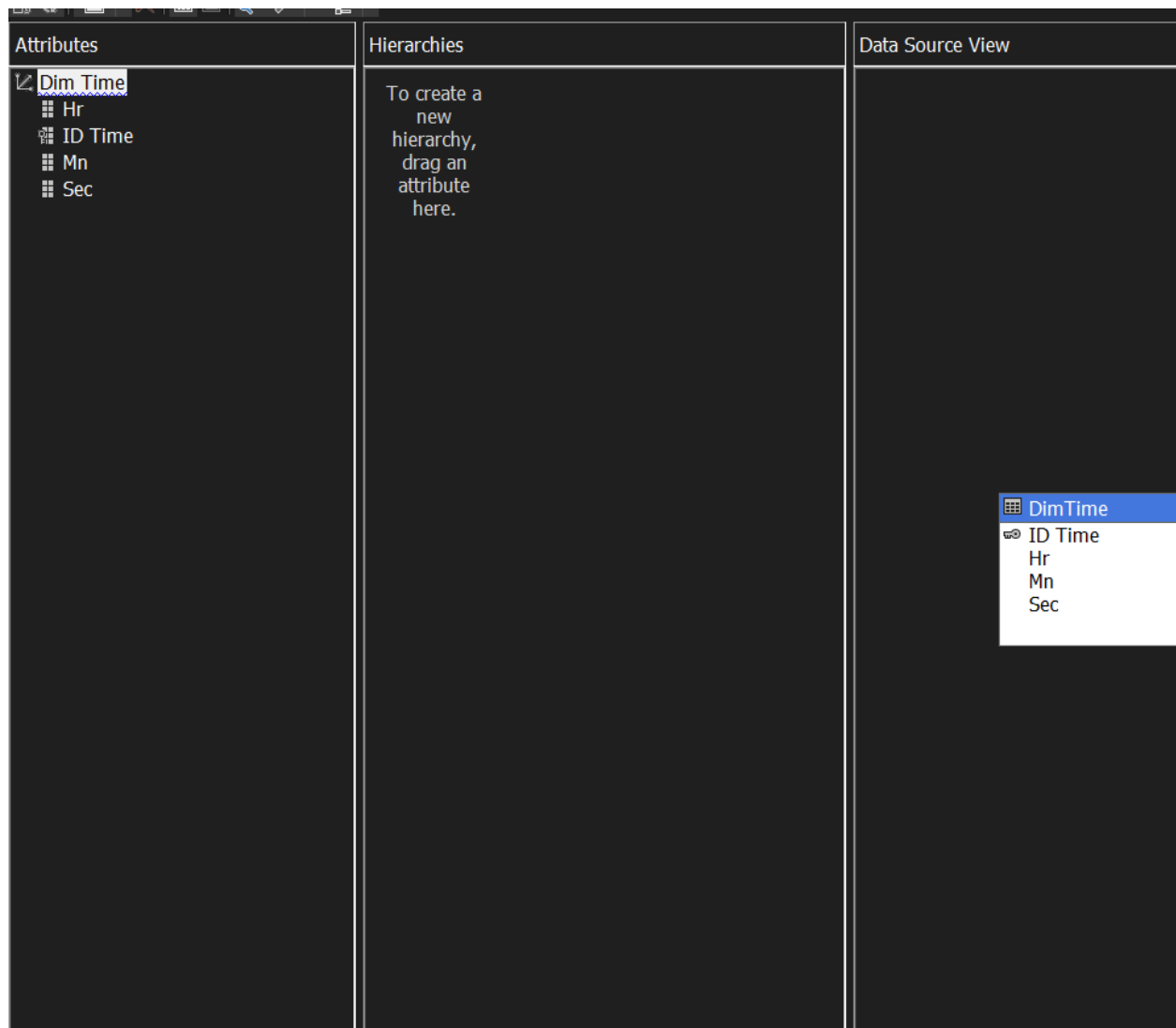
5.4.3 DimLocation

Attributes	Hierarchies	Data Source View
⬈ **Dim Location** ⫶ Country ⊞ ID Location ⫶ Latitude ⫶ Location Name ⫶ Longitude ⫶ Maximum Water Height m ⫶ Number Of Runups ⫶ Tsunami Intensity ⫶ Tsunami Magnitude Abe ⫶ Tsunami Magnitude Iida	To create a new hierarchy, drag an attribute here.	▦ DimLocation ⌐ø ID Location Country Location Name Latitude Longitude Maximum Water Height (m) Number of Runups Tsunami Magnitude (Abe) Tsunami Magnitude (Iida) Tsunami Intensity

5.4.4 DimDescription

Attributes	Hierarchies	Data Source View
Dim Description	To create a new hierarchy, drag an attribute here.	
Deposits		
Earthquake Magnitude		
ID Description		
More Info		
Tsunami Cause Code		
Tsunami Event Validity		
Vol		

DimDescription
ID Description
Tsunami Event Validity
Tsunami Cause Code
Earthquake Magnitude
Vol
More Info
Deposits

5.4.5 DimDetailDamage

Attributes	Hierarchies	Data Source View
Dim Detail Damage	To create a new hierarchy, drag an attribute here.	

Attributes:
- Damage Mil
- Damage Description
- Death Description
- Deaths
- Houses Damaged
- Houses Damaged Description
- Houses Destroyed
- Houses Destroyed Description
- ID Detail Damage
- Injuries
- Injuries Description
- Missing
- Missing Description

DimDetailDamage
- ID Detail Damage
- Deaths
- Death Description
- Missing
- Missing Description
- Injuries
- Injuries Description
- Damage ($Mil)
- Damage Description
- Houses Destroyed
- Houses Destroyed Description
- Houses Damaged
- Houses Damaged Description

5.4.6 DimTotalDamage

Attributes	Hierarchies	Data Source View

Attributes

- Dim Total Damage
- ID Total Damage
- Total Damage Mil
- Total Damage Description
- Total Death Description
- Total Deaths
- Total Houses Damaged
- Total Houses Damaged Description
- Total Houses Destroyed
- Total Houses Destroyed Description
- Total Injuries
- Total Injuries Description
- Total Missing
- Total Missing Description

Hierarchies

To create a new hierarchy, drag an attribute here.

Data Source View

- DimTotalDamage
- ID Total Damage
- Total Deaths
- Total Death Description
- Total Missing
- Total Missing Description
- Total Injuries
- Total Injuries Description
- Total Damage ($Mil)
- Total Damage Description
- Total Houses Destroyed
- Total Houses Destroyed Description
- Total Houses Damaged
- Total Houses Damaged Description

5.4.7 Run the Process for Deployment
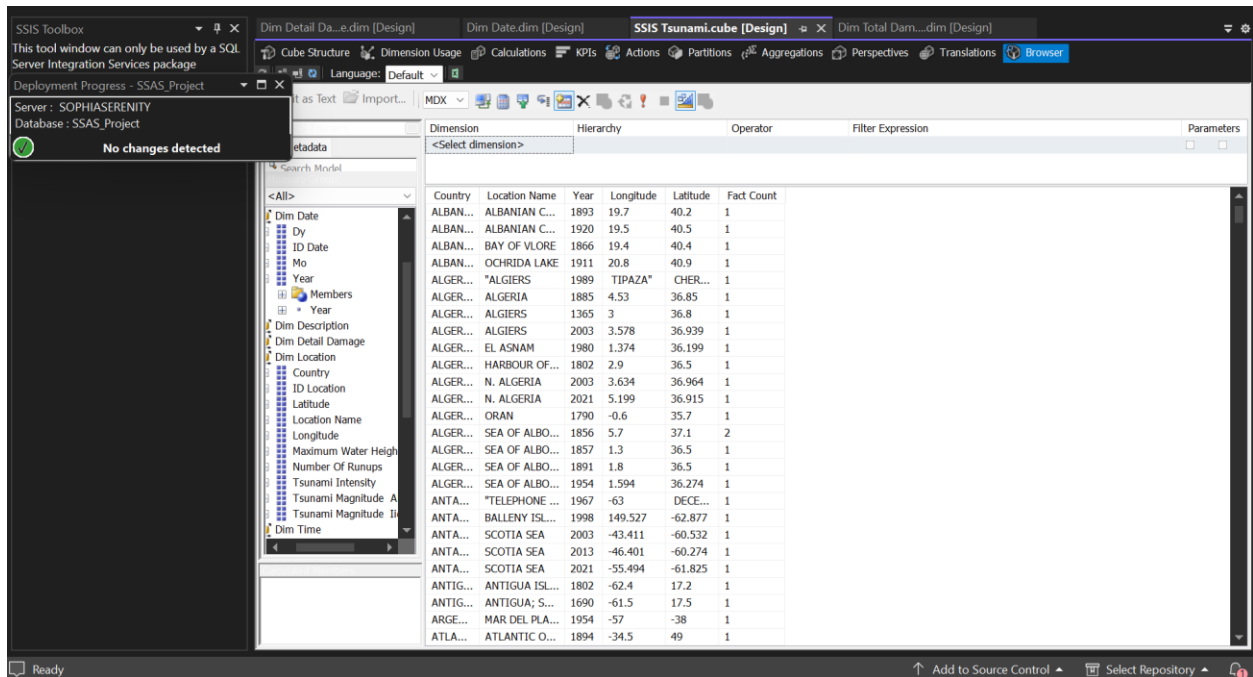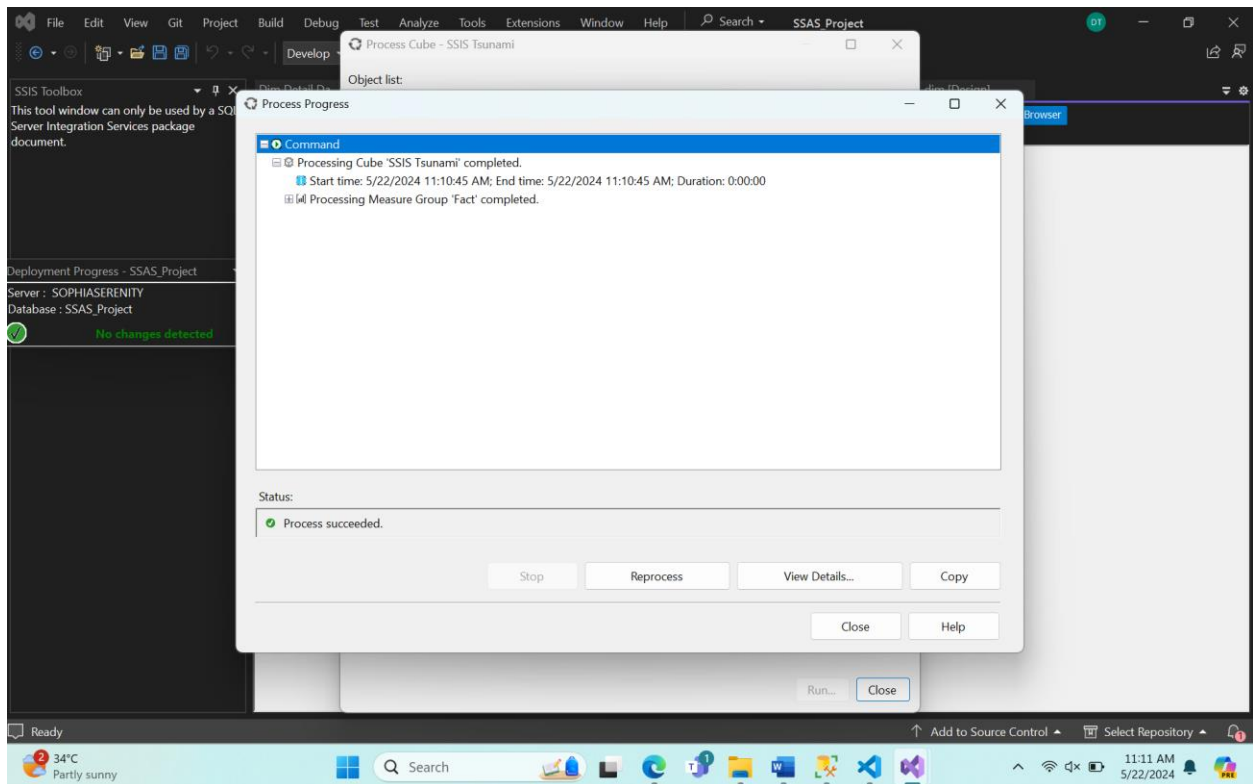
6　Query MDX

6.1　Query Tsunami year and description

6.1.1 SQL Query

```
-- 1. Query Tsunami date and description

SELECT
  {
    [Measures].[Fact Count]
  } ON COLUMNS,
  NON EMPTY
  (
    [Dim Date].[Year].[Year].Members *
    [Dim Description].[Tsunami Event Validity].[Tsunami Event Validity].Members *
    [Dim Description].[Tsunami Cause Code].[Tsunami Cause Code].Members *
    [Dim Description].[Earthquake Magnitude].[Earthquake Magnitude].Members *
    [Dim Description].[Vol].[Vol].Members
  ) ON ROWS
FROM [SSIS Tsunami]
```

6.1.2 Get SQL result

					Fact Count
1008	2	1	6.5		1
1026	3	1	7.5		1
103	3	1	7		1
1033	4	1			1
1039	1	1	6.8		1
1050	3	6		4845	1
1065	3	1			1
1068	3	1			1
1076	1	0			1
1088	3	1			1
1088	4	1	7		1
1092	-1	9			1
1096	4	1	8.4		1
1099	1	0			1
1099	4	1	8.4		1

6.2 Query Number of Deaths and their description
6.2.1 SQL Query

```
-- 2. Query number of deaths and their description

SELECT
  {
    [Measures].[Fact Count]
  } ON COLUMNS,
  NON EMPTY
  (
    [Dim Detail Damage].[Deaths].[Deaths].Members *
    [Dim Detail Damage].[Death Description].[Death Description].Members
  ) ON ROWS
FROM [SSIS Tsunami]
```

6.2.2 Get SQL Result

		Fact Count
11	1	2
110	3	1
1105	4	1
1169	4	1
12	1	5
1200	4	1
1223	4	1
1239	4	1
128	3	1
13	1	2
1300	4	1
13486	4	1
1362	4	1
14	1	2

6.3 Query Tsunami intensity over the year
6.3.1 SQL query

```
-- 4. Query total of deaths over the year with country

SELECT
  NON EMPTY {
    [Measures].[Fact Count]
  } ON COLUMNS,
  NON EMPTY {
    [Dim Location].[Country].[Country].Members *
    [Dim Date].[Year].[Year].Members *
    [Dim Total Damage].[Total Deaths].[Total Deaths].Members
  } ON ROWS
FROM [SSIS Tsunami]
```

6.3.2 Get SQL Result

			Fact Count
"MICRONESIA	1837		1
"MICRONESIA	1912		2
"MICRONESIA	1925		1
ALBANIA	1273		1
ALBANIA	1833		1
ALBANIA	1851	2000	1
ALBANIA	1866		3
ALBANIA	1866	60	2
ALBANIA	1893		1
ALBANIA	1911		1
ALBANIA	1920		1
ALBANIA	346		1
ALBANIA	-58		1
ALGERIA	1365		1
ALGERIA	1790	3000	1
ALGERIA	1802		1
ALGERIA	1856	5	1
ALGERIA	1856	8	1
ALGERIA	1857		1
ALGERIA	1885		1
ALGERIA	1891	26	1

6.4 Query total of deaths over the year with country
6.4.1 SQL Query

```
-- 4. Query total of deaths over the year with country

SELECT
  NON EMPTY {
    [Measures].[Fact Count]
  } ON COLUMNS,
  NON EMPTY {
    [Dim Location].[Country].[Country].Members *
    [Dim Date].[Year].[Year].Members *
    [Dim Total Damage].[Total Deaths].[Total Deaths].Members
  } ON ROWS
FROM [SSIS Tsunami]
```

6.4.2 Get SQL Result

			Fact Count
"MICRONESIA	1837		1
"MICRONESIA	1912		2
"MICRONESIA	1925		1
ALBANIA	1273		1
ALBANIA	1833		1
ALBANIA	1851	2000	1
ALBANIA	1866		3
ALBANIA	1866	60	2
ALBANIA	1893		1
ALBANIA	1911		1
ALBANIA	1920		1
ALBANIA	346		1
ALBANIA	-58		1
ALGERIA	1365		1
ALGERIA	1790	3000	1
ALGERIA	1802		1
ALGERIA	1856	5	1
ALGERIA	1856	8	1
ALGERIA	1857		1
ALGERIA	1885		1
ALGERIA	1891	36	1
ALGERIA	1954	1243	1

6.5 Query number of houses damaged and their descriptions

6.5.1 SQL Query

```
-- 5.  Query for number of houses damaged and their descriptions

SELECT
  {
    [Measures].[Fact Count]
  } ON COLUMNS,
  NON EMPTY
  (
    [Dim Detail Damage].[Houses Damaged].[Houses Damaged].Members *
    [Dim Detail Damage].[Houses Damaged Description].[Houses Damaged Description].Membe
  ) ON ROWS
FROM [SSIS Tsunami]
```

6.5.2 Get SQL Result

		Fact Count
182	3	1
19	1	1
2		6
22	1	1
238	3	1
280920	4	1
2844	4	1
3		9
3	1	1
300	3	1
39191	4	1
4		5
478	3	1
500	3	1

6.6 Query for the total number of tsunamis by time of day with descriptions

6.6.1 SQL Query

```
-- 6. Query for the total number of tsunamis by time of day with descriptions

SELECT
  {
    [Measures].[Fact Count]
  } ON COLUMNS,
  NON EMPTY
  (
    [Dim Time].[Hr].[Hr].Members *
    [Dim Description].[Tsunami Event Validity].[Tsunami Event Validity].Members *
    [Dim Description].[Tsunami Cause Code].[Tsunami Cause Code].Members
  ) ON ROWS
FROM [SSIS Tsunami]
```

6.6.2 Get SQL Result

			Fact Count
1	1	9	1
1	-1	1	1
1	2	1	8
1	2	3	1
1	3	1	11
1	3	3	1
1	3	9	1
1	4	1	27
1	4	3	2
1	4	6	1
1	4	9	1
10	0	1	2
10	0	2	1
10	1	0	1

6.7 Query Countries and location has tsunami over the months

6.7.1 SQL Query

```
-- 7. Countries and location has tsunami over the dim date

SELECT
  {
    [Measures].[Fact Count]
  } ON COLUMNS,
  NON EMPTY
  (
    [Dim Location].[Country].[Country].Members *
    [Dim Location].[Location Name].[Location Name].Members *
    [Dim Date].[Mo].[Mo].Members

  ) ON ROWS
FROM [SSIS Tsunami]
```

6.7.2 Get SQL Result

			Fact Count
"MICRONESIA	FED. STATES OF"	10	2
"MICRONESIA	FED. STATES OF"	12	2
ALBANIA	ADRIATIC SEA		1
ALBANIA	ALBANIA		1
ALBANIA	ALBANIA	1	1
ALBANIA	ALBANIAN COASTS	1	2
ALBANIA	ALBANIAN COASTS	10	1
ALBANIA	ALBANIAN COASTS	12	1
ALBANIA	ALBANIAN COASTS	3	2
ALBANIA	ALBANIAN COASTS	6	1
ALBANIA	ALBANIAN COASTS	9	1
ALBANIA	BAY OF VLORE	3	1
ALBANIA	OCHRIDA LAKE	2	1
ALGERIA	"ALGIERS	10	1
ALGERIA	ALGERIA	1	1
ALGERIA	ALGIERS	1	1
ALGERIA	ALGIERS	5	1
ALGERIA	EL ASNAM	10	1
ALGERIA	HARBOUR OF ALGIERS	11	1
ALGERIA	N. ALGERIA	3	1
ALGERIA	N. ALGERIA	5	1

6.8 Top 10 Countries by Event Count in 2020

6.8.1 SQL Query

```
--8. Top 10 Countries by Event Count in 2020
SELECT
  NON EMPTY {
    [Measures].[Fact Count]
  } ON COLUMNS,
  NON EMPTY {
    TOPCOUNT(
      [Dim Location].[Country].[Country].Members,
      10,
      [Measures].[Fact Count]
    )
  } ON ROWS
FROM [SSIS Tsunami]
WHERE ([Dim Date].[Year].&[2020])
```

6.8.2 Get SQL Result

	Fact Count
USA	4
CANADA	1
CHILE	1
CUBA	1
GREECE	1
INDONESIA	1
MEXICO	1
NEW ZEALAND	1
RUSSIA	1
SPAIN	1

6.9 Top 3 Years by Event Count
6.9.1 SQL Query

```
--9. Top 3 Years by Event Count
SELECT
  NON EMPTY {
    [Measures].[Fact Count]
  } ON COLUMNS,
  NON EMPTY {
    TOPCOUNT(
      [Dim Date].[Year].[Year].Members,
      3,
      [Measures].[Fact Count]
    )
  } ON ROWS
FROM [SSIS Tsunami]
```

6.9.2 Get SQL Result

	Fact Count
1928	25
1938	24
1868	23

6.10 Query for the total damage ($) by country and year
6.10.1 SQL Query

```
--10. Top 3 Countries by Event Count for Each Year

SELECT
  NON EMPTY {
    [Measures].[Fact Count]  -- Replace with the correct measure name for deaths
  } ON COLUMNS,
  NON EMPTY {
    GENERATE(
      [Dim Date].[Year].[Year].Members,
      TOPCOUNT(
        [Dim Location].[Country].[Country].Members,
        3,
        [Measures].[Fact Count]  -- Replace with the correct measure name for deaths
      )
    )|
  } ON ROWS
FROM [SSIS Tsunami]
```

6.10.2 Get SQL Result

	Fact Count
IRAN	6
	1
"MARSHALL ISLANDS	1
JAPAN	371
UKRAINE	12
SYRIA	8
TURKEY	67
GREECE	179
ISRAEL	9
CHINA	43
UK	23
ITALY	160
NORTH SEA	1
CYPRUS	5
SOUTH KOREA	14
ALBANIA	13
ALGERIA	14

7   Report Data (by using PowerBI)

7.1 Sum of Deaths by Year
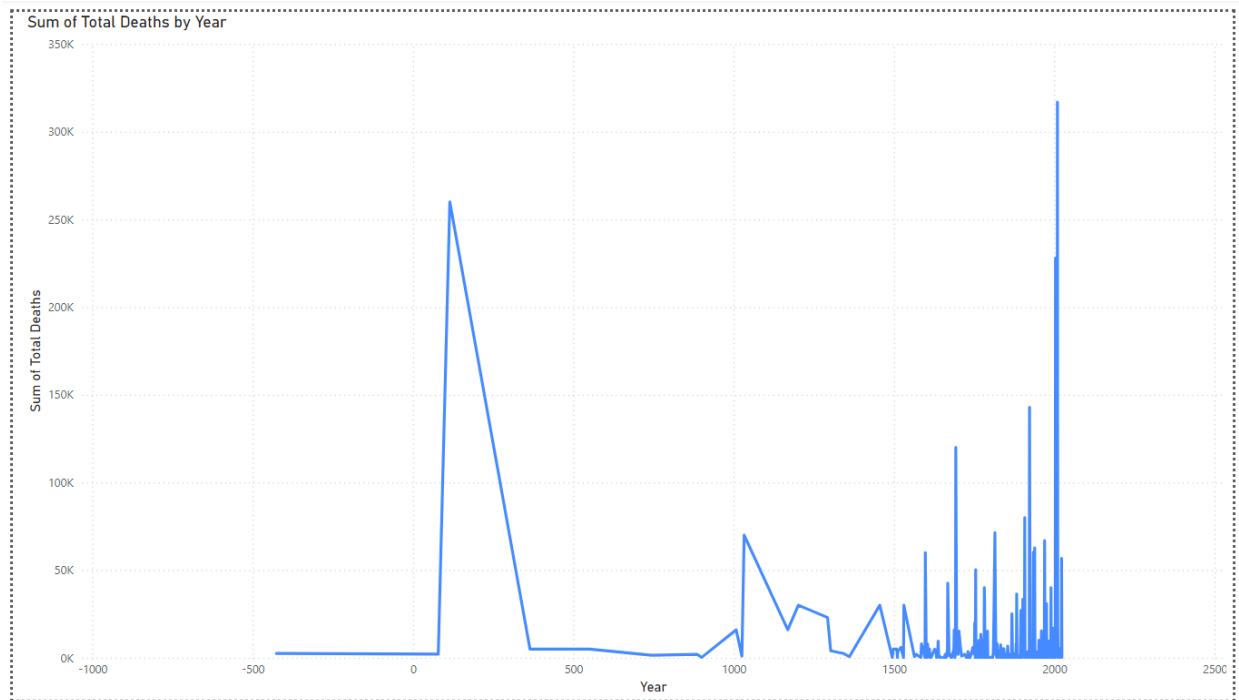
7.2 Sum of Tsunami Intensity by Year

7.3 Count of Total Houses Damaged by Year
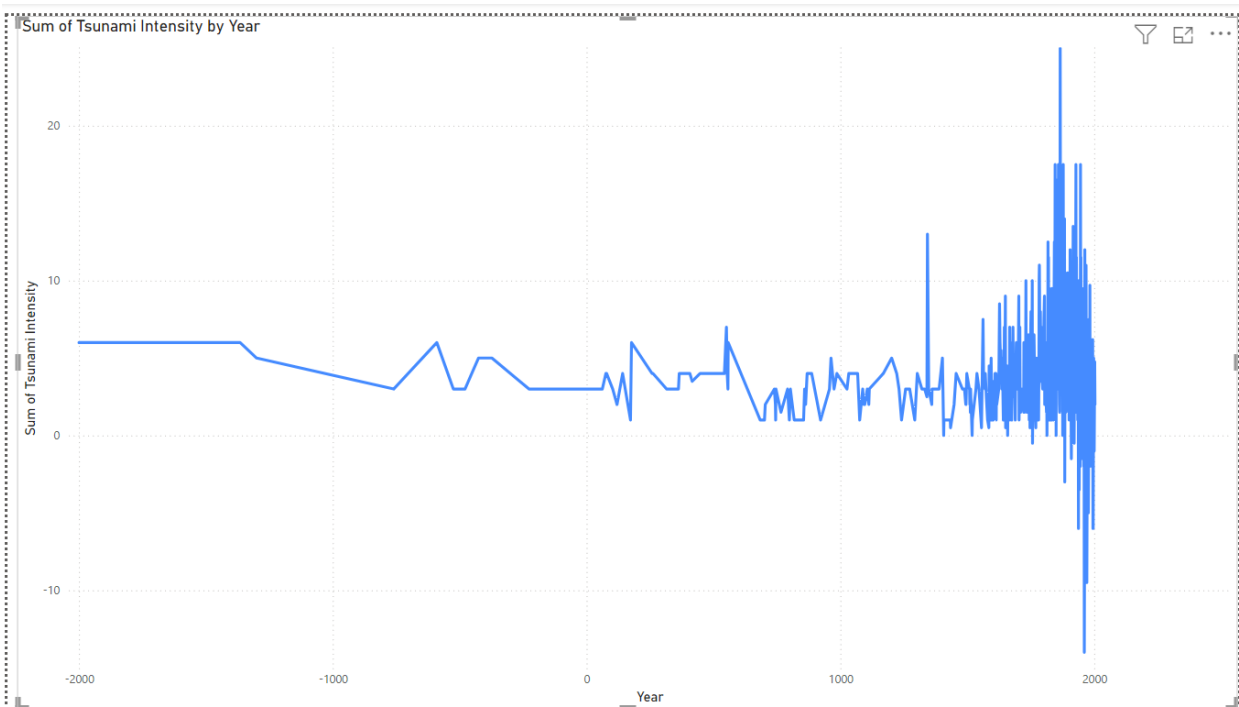
Count of Total Houses Damaged by Year

7.4 Sum of Maximum Water Height (m) by Year
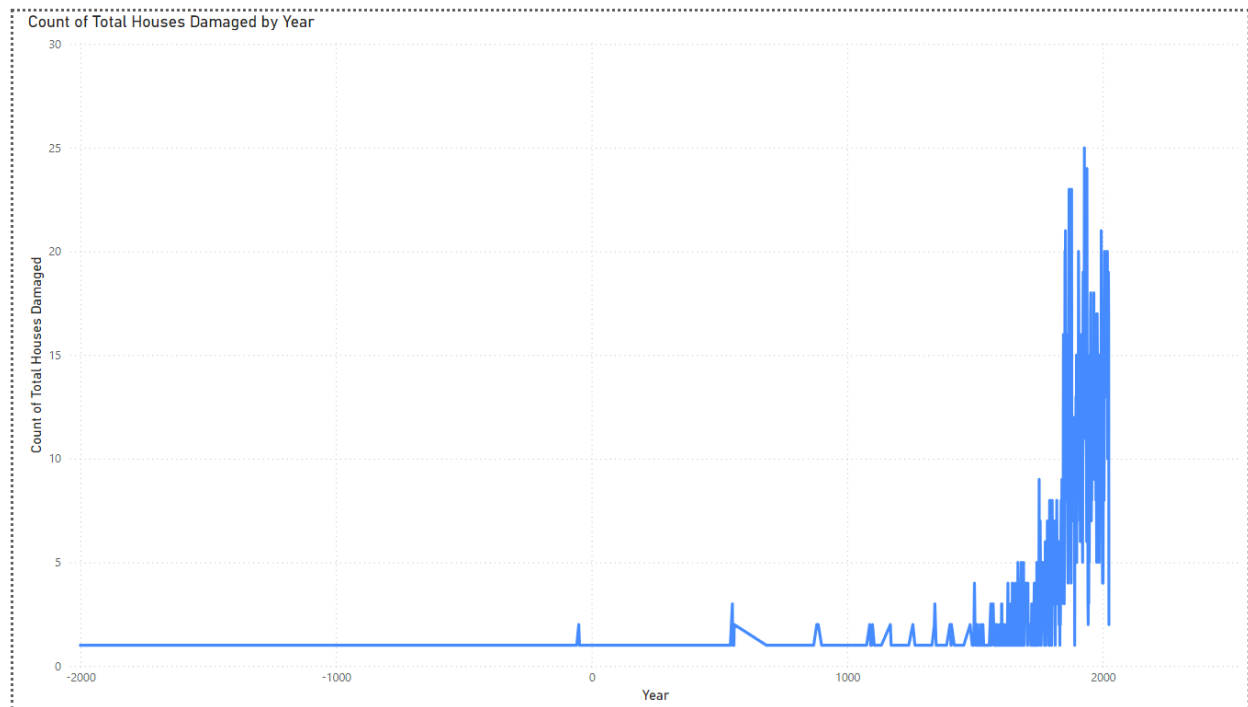
Sum of Maximum Water Height (m) by Year
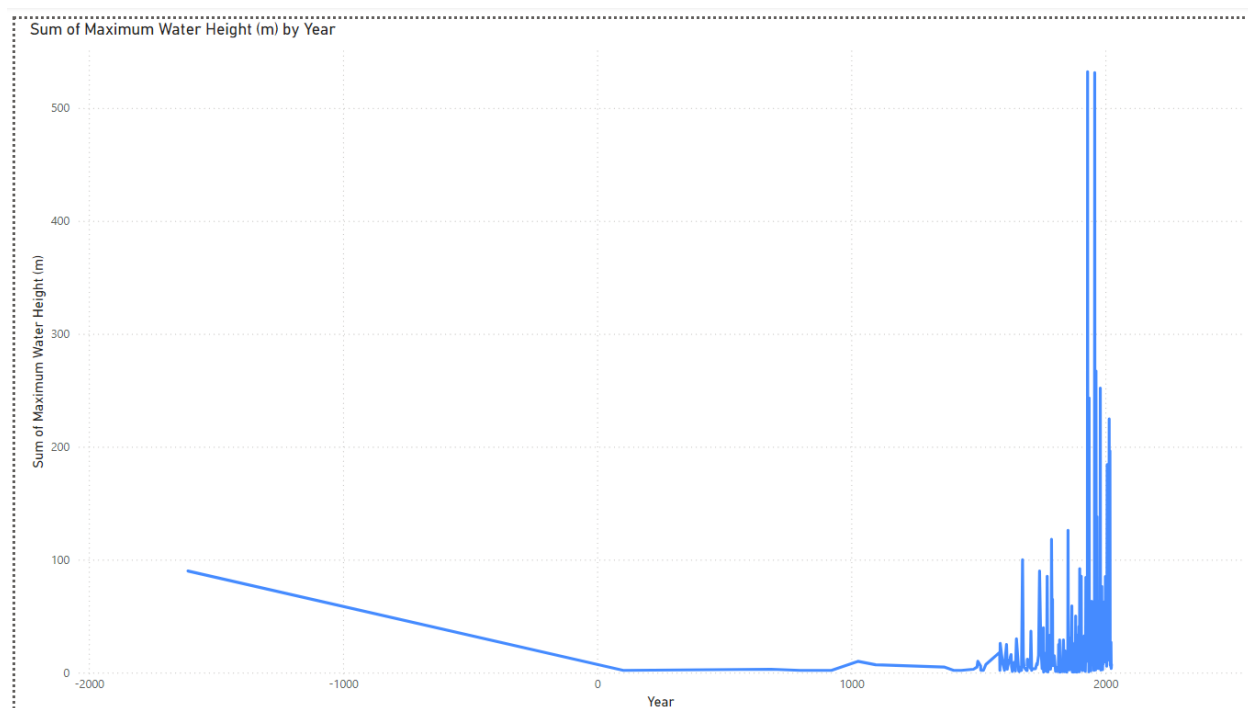
7.5 Sum of Vol by Year

Sum of Vol by Year

7.6 Sum of Earthquake Magnitude by Year
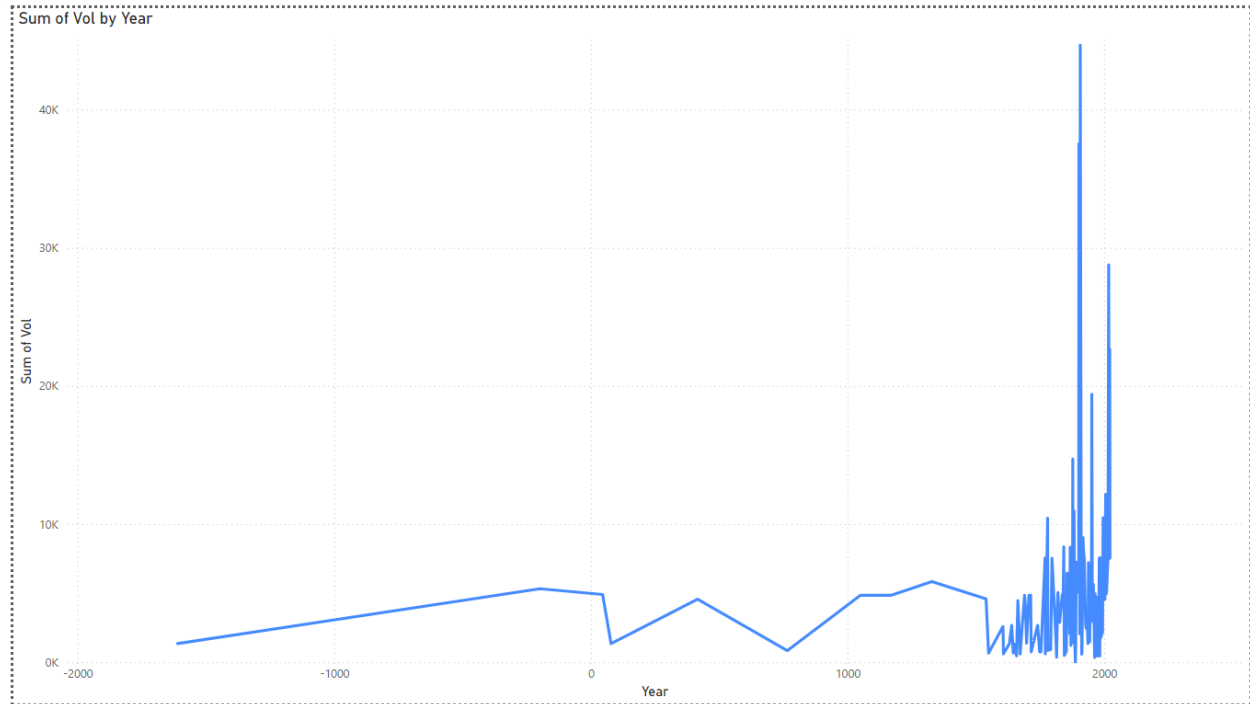
Sum of Earthquake Magnitude by Year

7.7 Sum of Number of Runups and Sum of Earthquake Magnitude by Year
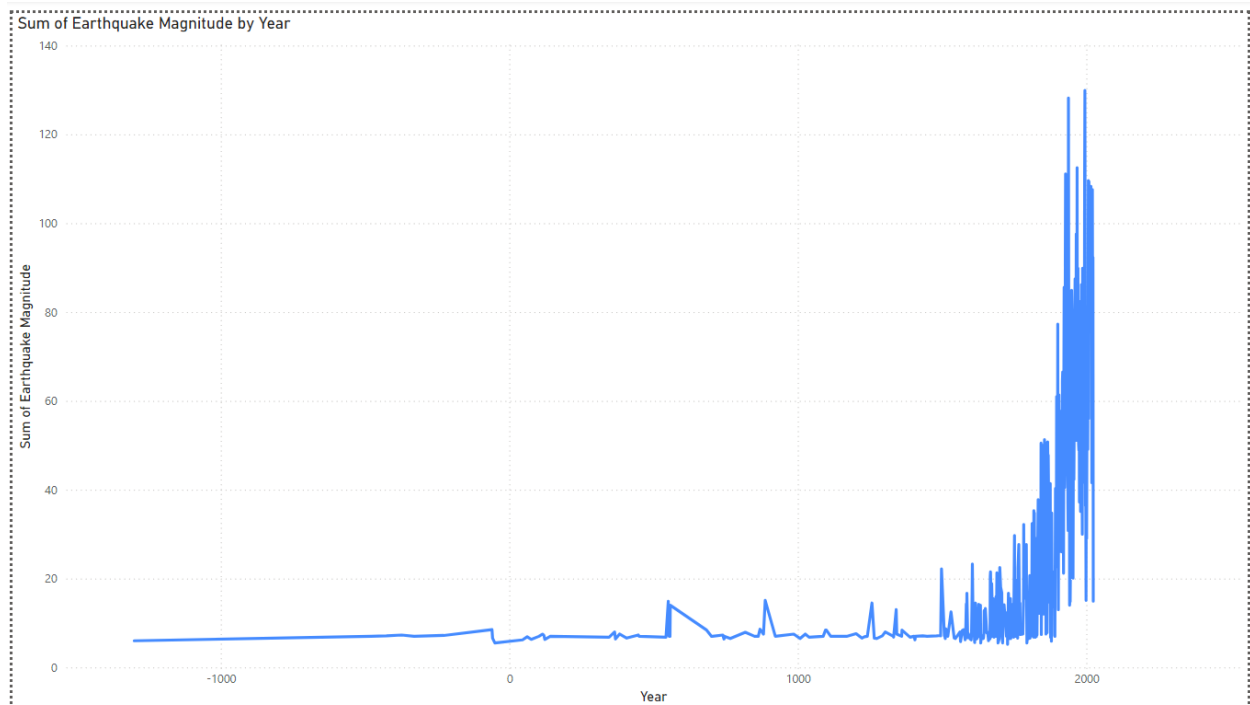
Sum of Number of Runups and Sum of Earthquake Magnitude by Year

7.8 Count of Total Damage ($Mil) by Country

Count of Total Damage ($Mil) by Country

7.9 Sum of Tsunami Intensity in Month by Country

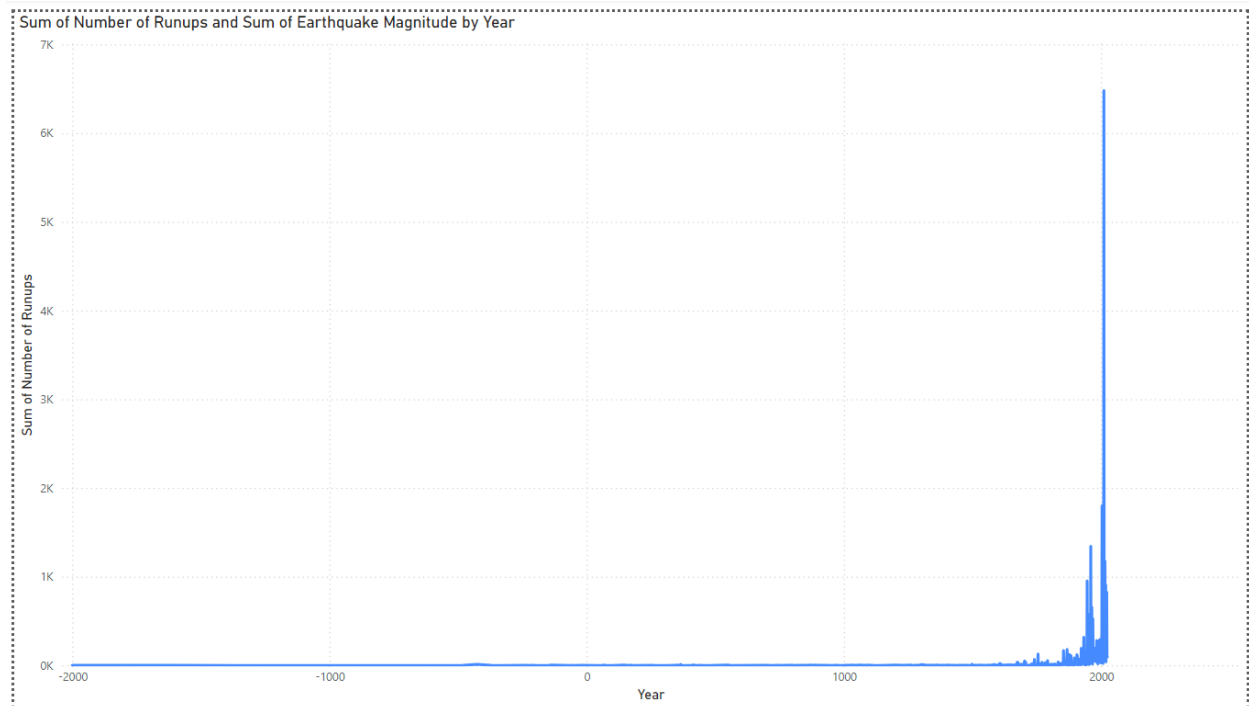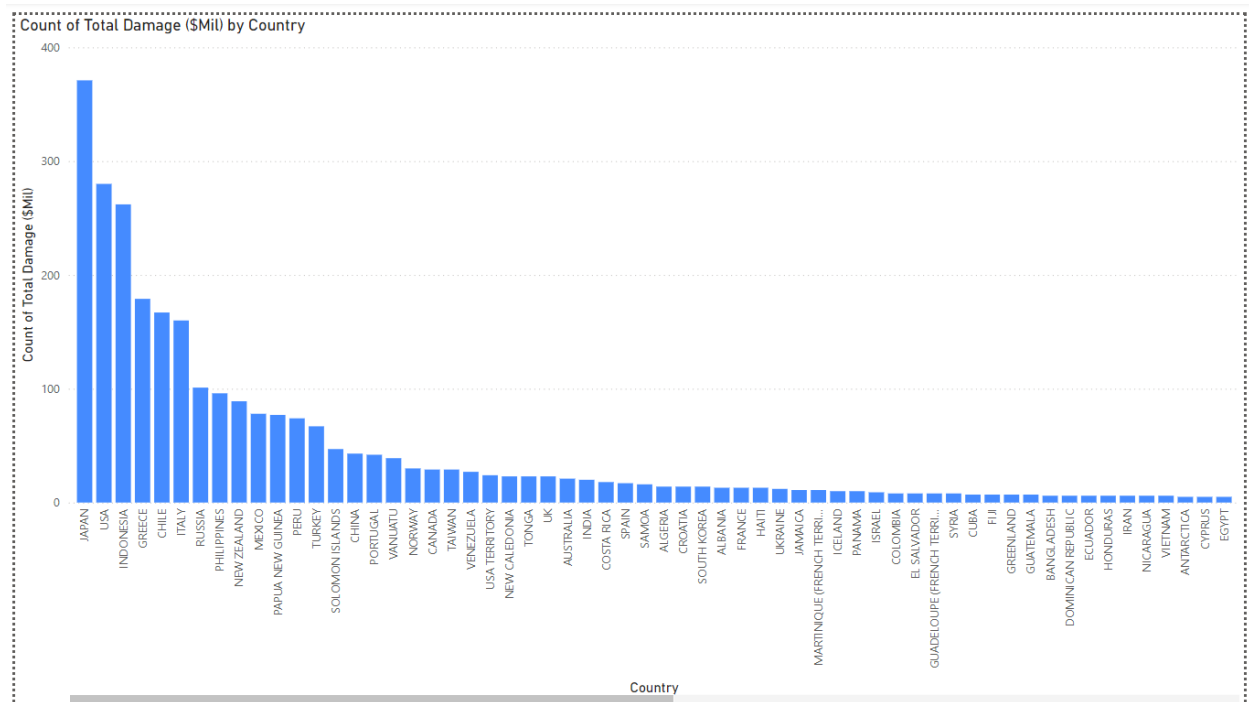| Country		1	2	3	4	5	6	7	8	9	10	11	12	Total
ALBANIA	3.00	11.00		10.00		3.00			3.00	3.00			5.00	**38.00**
ALGERIA						1.50		8.00		6.00	2.00			**17.50**
ANTARCTICA													0.00	**0.00**
ATLANTIC OCEAN											2.00			**2.00**
AUSTRALIA									-5.00					**-5.00**
BULGARIA				4.00										**4.00**
CANADA			0.00			0.00	2.00							**2.00**
CHILE	2.00	1.50	7.50	9.00	4.50	13.50	0.00	7.50	24.00	0.50	6.20	17.00	7.50	**100.70**
CHINA		1.00	-1.00			2.00	1.00		1.50	2.00	1.00	2.00		**9.50**
COLOMBIA						1.00							2.50	**3.50**
COOK ISLANDS				2.50										**2.50**
COSTA RICA						-3.00			3.00		-2.00		-7.00	**-9.00**
CROATIA		3.00		5.00	4.00				6.00	10.00	4.00		3.00	**35.00**
CYPRUS		2.00				4.00	4.00			2.00				**12.00**
DEAD SEA	3.00													**3.00**
EAST CHINA SEA									0.00					**0.00**
ECUADOR		5.00	0.50								1.50			**7.00**
EGYPT	3.00					3.00								**6.00**
EL SALVADOR			2.00						1.50	0.50			1.50	**5.50**
FIJI									1.00	1.00				**2.00**
FRANCE		4.00				3.00			4.00					**11.00**
FRENCH POLYNESIA									0.50					**0.50**
GREECE	34.00	17.00	36.00	13.00	28.00	12.00	24.00	22.00	26.00	16.00	14.00	24.00	16.00	**282.00**
GUATEMALA										-1.00				**-1.00**
INDONESIA	5.00	14.80	16.30	18.50	9.50	10.00	9.50	4.00	20.00	9.00	9.82	15.50	15.20	**157.12**

7.10 Count of Tsunami Days in Months

Mo		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	Total
	151																																**151**
1	13	14	5	7	8	4	5	4	11	10	6	4	5	7	8	10	10	4	7	8	9	7	7	8	4	9	5	7	8	6	7	11	**238**
2	4	2	12	8	7	5	8	10	8	8	6	11	3	10	7	3	8	8	3	6	7	8	5	10	5	5	6	5	4	4			**196**
3	13	7	7	4	10	5	5	5	4	11	2	10	9	4	8	5	8	8	11	9	12	5	8	6	7	9	6	3	8	3	7	12	**231**
4	7	8	6	8	7	7	6	5	6	7	9	7	7	11	7	4	8	6	7	9	6	7	8	7	9	3	3	5	5	4	7		**206**
5	13	6	6	5	11	8	9	16	8	1	6	6	5	7	8	6	6	7	3	10	10	6	8	10	6	9	9	9	6	4	6	5	**234**
6	4	4	4	8	2	5	2	4	6	5	7	4	10	11	4	13	8	5	12	4	6	5	6	7	6	4	9	7	5	8	3		**188**
7	8	6	4	6	5	5	4	3	8	13	12	6	9	6	3	5	5	8	11	7	10	6	3	7	6	11	6	3	6	8	11	5	**216**
8	14	7	11	4	7	6	10	5	11	12	9	11	11	6	9	11	11	9	8	11	11	10	5	9	6	5	5	9	3	9	4	6	**265**
9	10	12	8	4	6	12	7	7	14	7	6	7	7	7	7	4	13	3	6	6	9	10	4	6	10	9	6	6	6	6	8		**233**
10	12	1	7	7	10	8	10	5	7	12	9	8	10	5	5	2	7	4	14	9	13	5	10	3	4	7	10	3	7	5	4	11	**234**
11	7	13	8	1	10	10	7	6	5	4	7	12	5	11	10	7	11	10	11	9	5	11	12	6	11	9	5	4	11	9	9		**256**
12	8	7	5	6	15	4	4	6	4	11	3	4	6	5	5	4	9	11	10	3	6	9	5	5	4	9	11	4	11	10	4	5	**213**
Total	**264**	**87**	**83**	**68**	**98**	**79**	**77**	**76**	**92**	**101**	**82**	**90**	**87**	**90**	**81**	**74**	**104**	**83**	**103**	**91**	**104**	**89**	**81**	**84**	**78**	**89**	**81**	**62**	**78**	**78**	**69**	**58**	**2861**

8 Pivot Table

8.1 Dataset

8.2 Statistics on the number of deaths and causes over the years

8.2.1 Top 10 Number of Tsunami Event Validity per Country
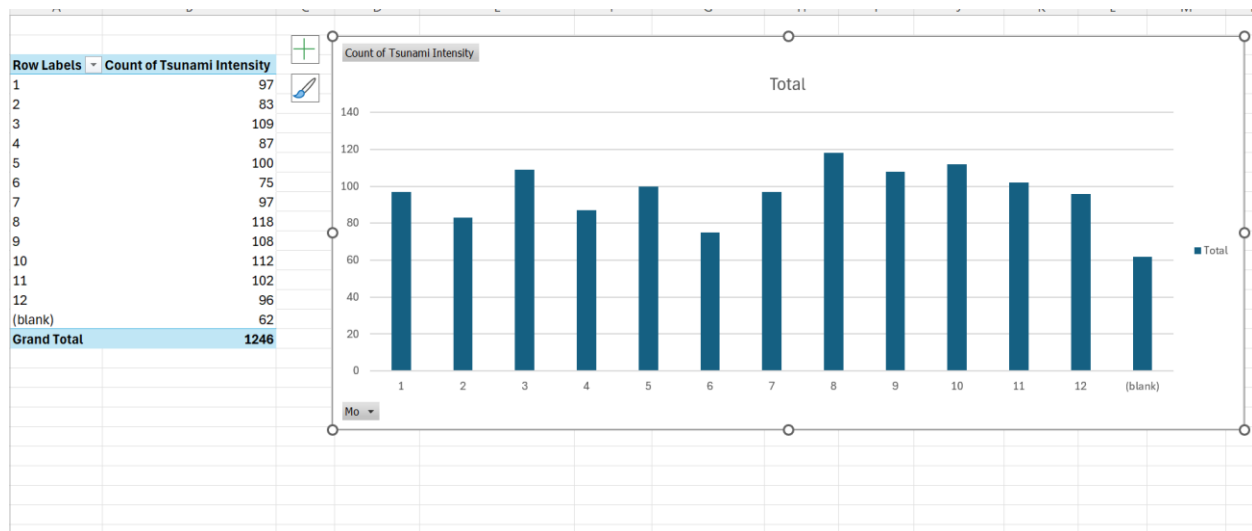
Row Labels	Count of Tsunami Event Validity
CHILE	167
GREECE	179
INDONESIA	262
ITALY	160
JAPAN	371
MEXICO	78
NEW ZEALAND	89
PHILIPPINES	96
RUSSIA	101
USA	280
Grand Total	**1783**

8.2.2 Number of Earthquake magnitude over the year

Row Labels	Count of Earthquake Magnitude
-2000	
-1610	
-1365	
-1300	1
-760	
-590	
-525	
-480	
-479	1
-426	1
-373	1
-330	1
-326	
-227	1
-226	
-223	
-216	
-210	
-198	
-150	
-138	
-60	1
-58	1
-50	1
-47	

Count of Earthquake Magnitude
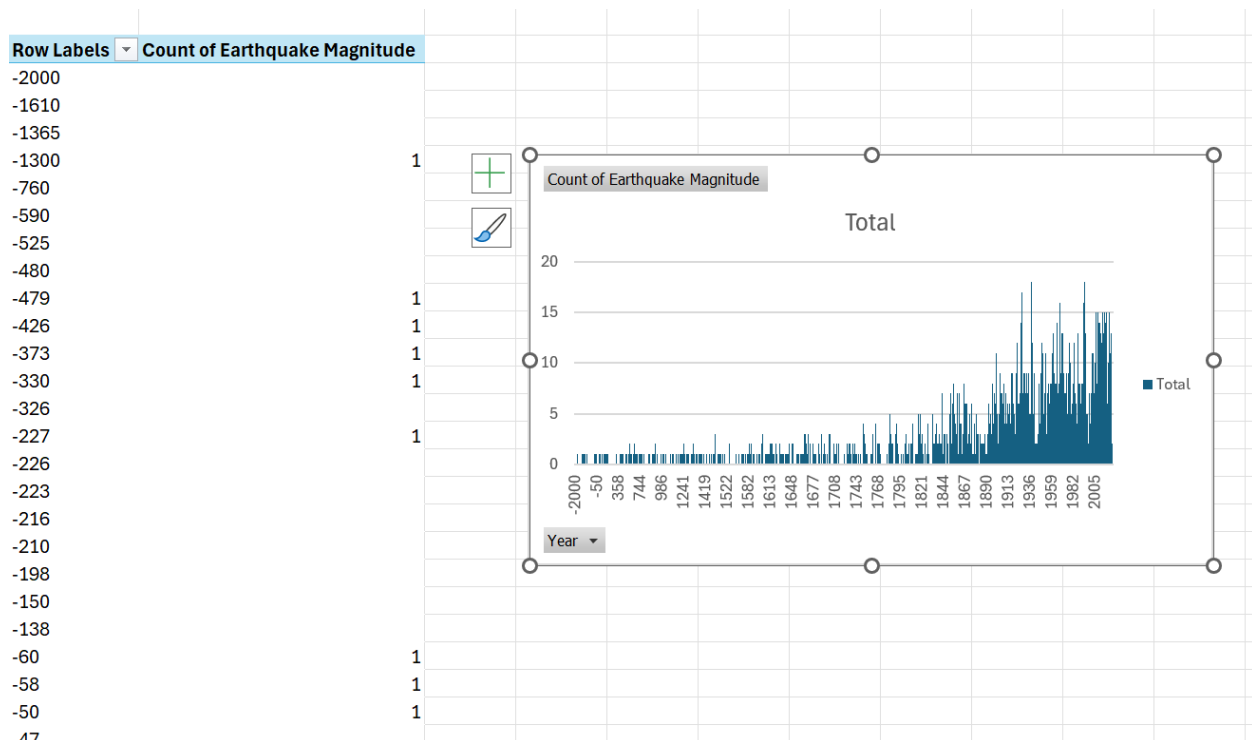
8.2.3 Number of Tsunami Intensity over the month

Row Labels	Count of Tsunami Intensity
1	97
2	83
3	109
4	87
5	100
6	75
7	97
8	118
9	108
10	112
11	102
12	96
(blank)	62
Grand Total	**1246**

Count of Tsunami Intensity

8.2.4 Top 10 of Country has Damage ($Mil)

Row Labels	Sum of Damage Description
CHILE	64
GREECE	49
INDONESIA	145
ITALY	29
JAPAN	225
NORWAY	26
PAPUA NEW GUINEA	38
PERU	42
PHILIPPINES	41
USA	82
Grand Total	741

Sum of Damage Description
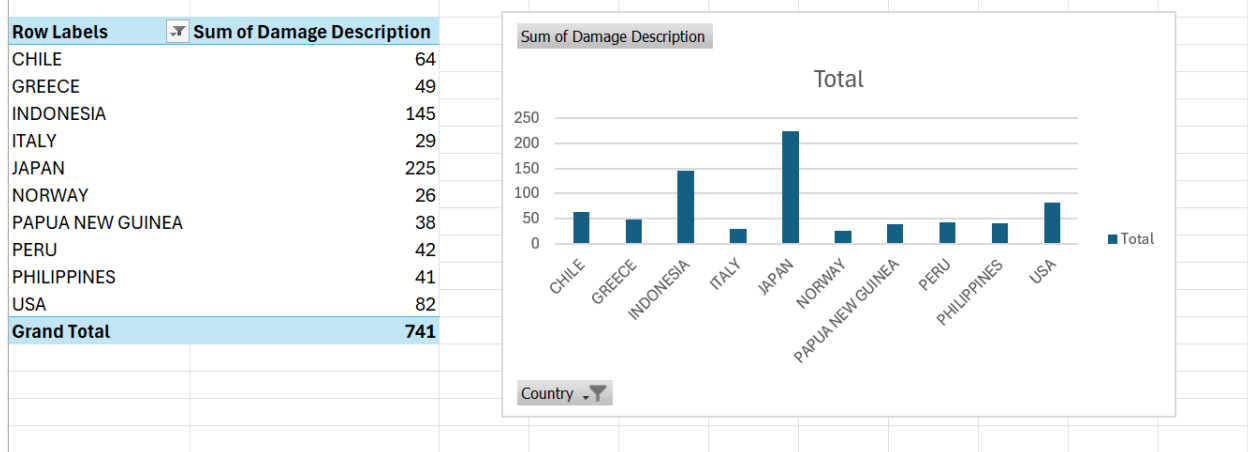
Total

8.2.5 Sum of Total House damage description

Row Labels	Sum of Total Houses Damaged Description
-2000	
-1610	
-1365	
-1300	
-760	
-590	
-525	
-480	
-479	
-426	
-373	
-330	
-326	
-227	
-226	
-223	
-216	
-210	
-198	
-150	
-138	
-60	
-58	
-50	
-47	

Sum of Total Houses Damaged Description
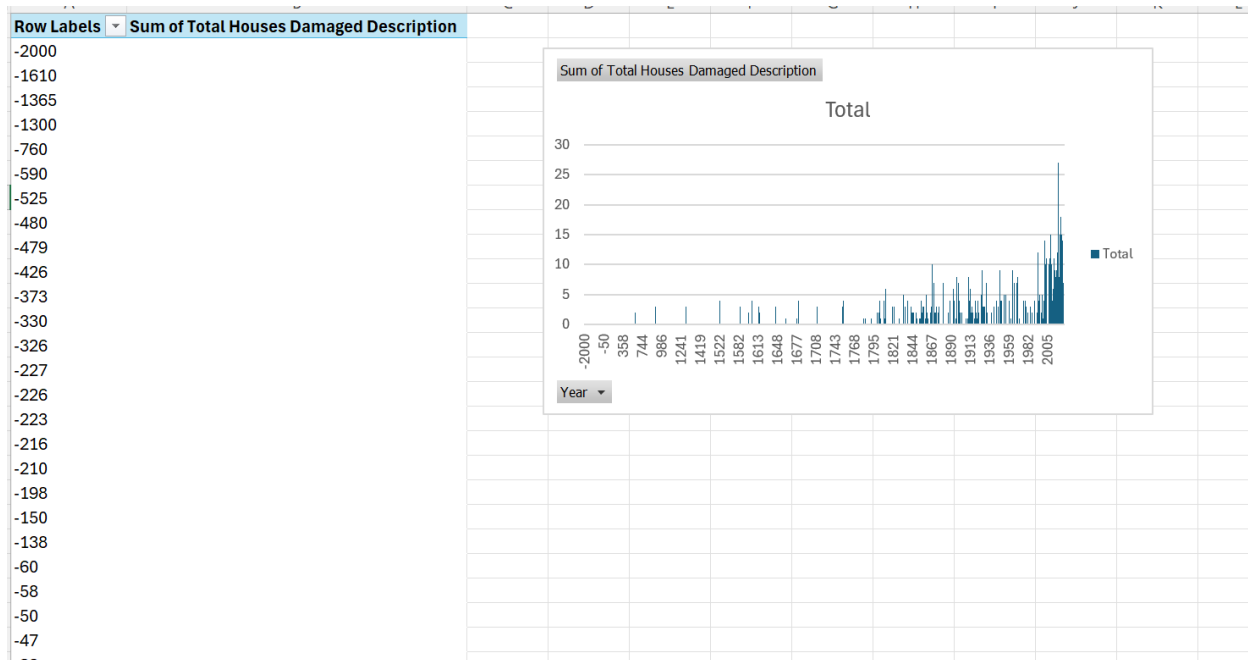
Total

8.2.6 Count of Country per Tsunami Cause

Column Labels													
	0	1	2	3	4	5	6	7	8	9	10	11	(blank)	Grand Total
Count of Country	297	1981	28	128	19	2	135	14	127	122	2	2		3	2860
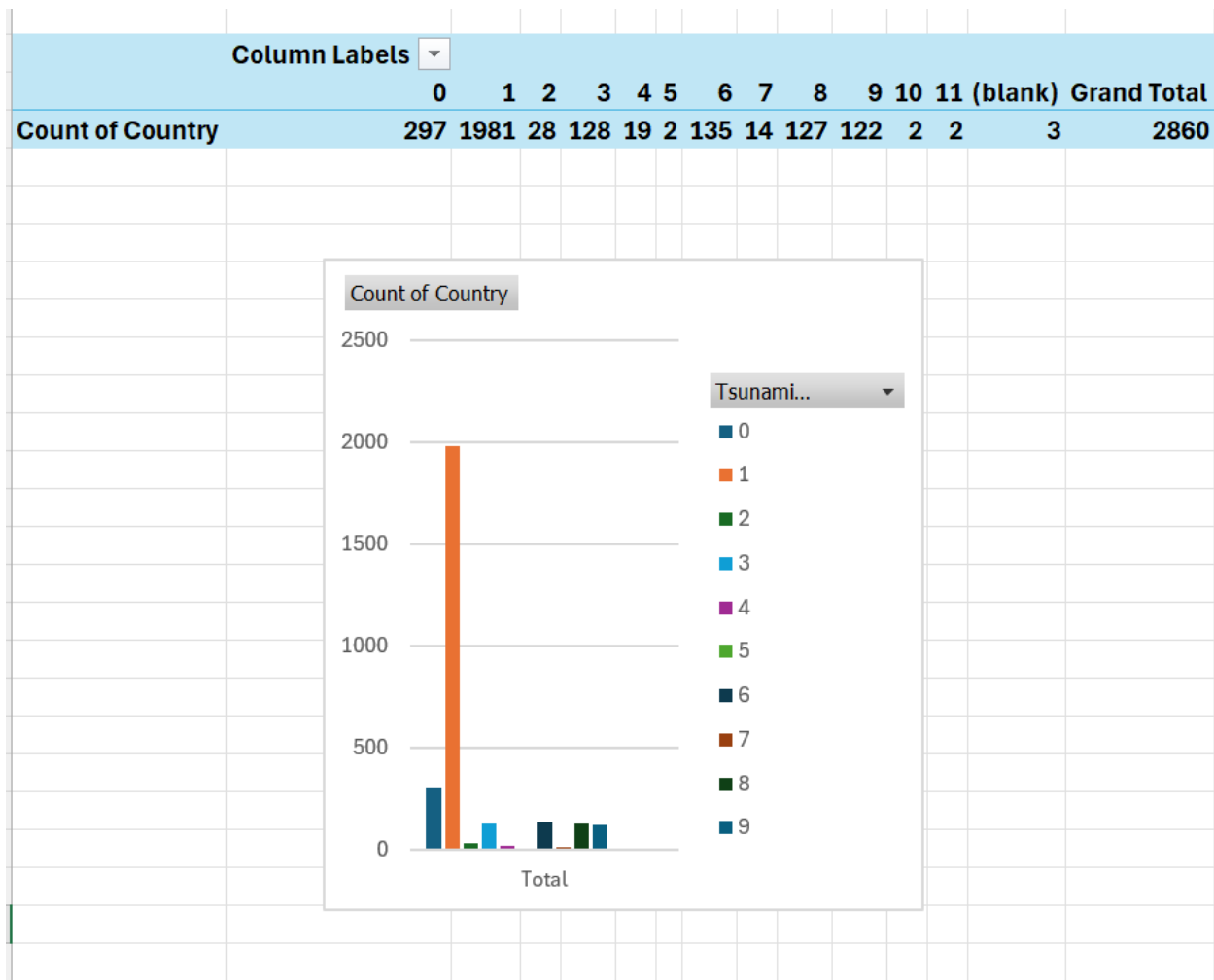
8.2.7 Top 10 of Country has Volcano

Row Labels	Sum of Vol
GREECE	22628
INDONESIA	80976
ITALY	86846
JAPAN	106636
MARTINIQUE (FRENCH TERRITORY)	33342
PAPUA NEW GUINEA	35895
RUSSIA	21461
SAMOA	52241
USA	58619
VANUATU	29007
Grand Total	527651
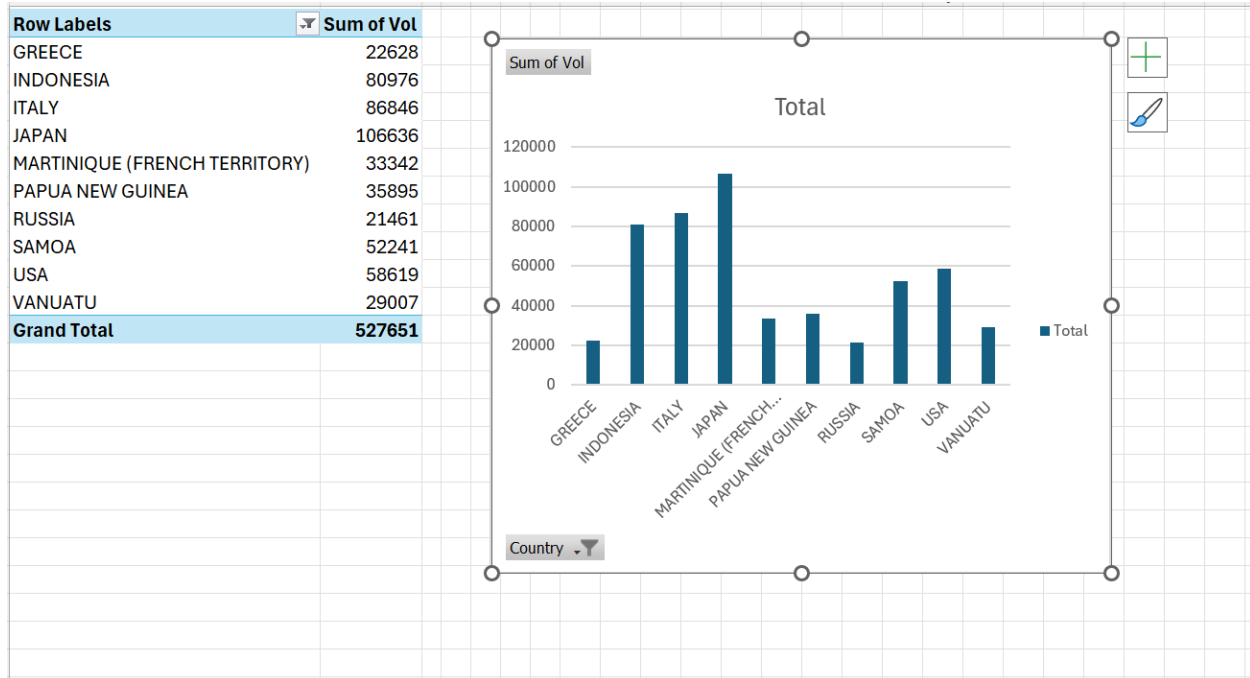
8.2.8 Top 10 of Country has total runnup highest

Row Labels	Sum of Number of Runups
CHILE	4043
INDONESIA	4203
JAPAN	11172
PAPUA NEW GUINEA	453
PERU	571
RUSSIA	1588
SAMOA	656
SOLOMON ISLANDS	625
TONGA	914
USA	2495
Grand Total	26720
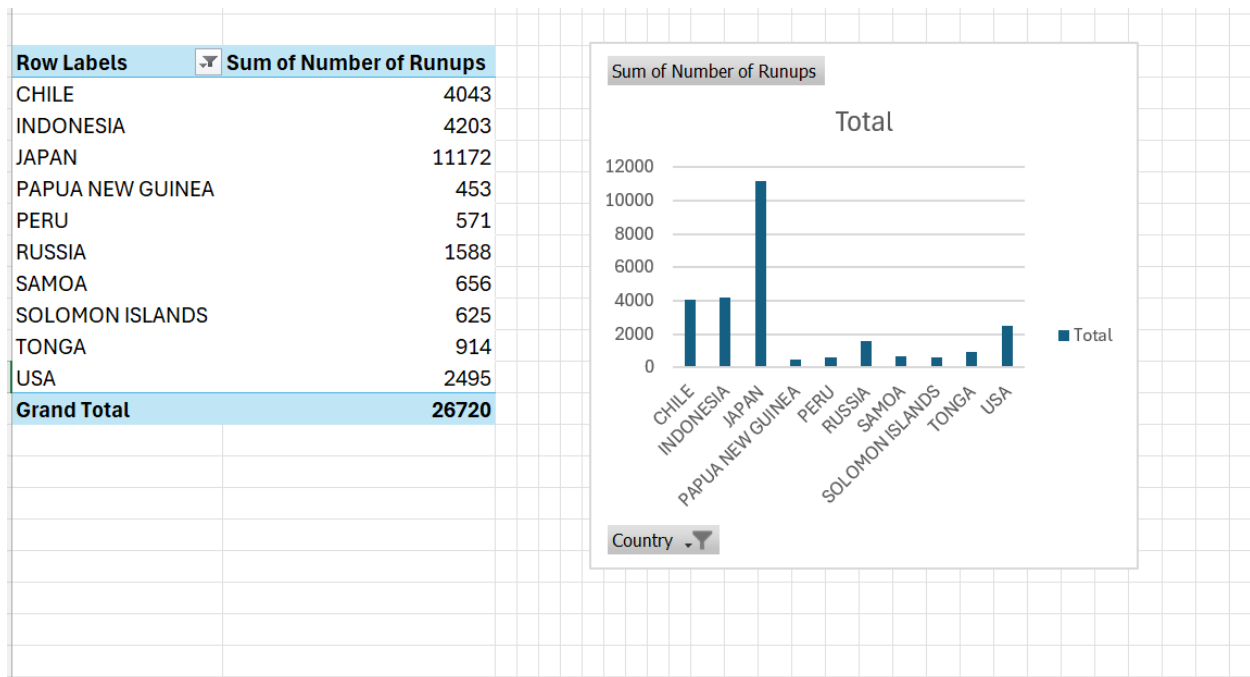
8.2.9 Top 10 of Year has number of deaths highest

Row Labels ⊤	Sum of Total Deaths
115	260000
1033	70000
1693	120000
1815	71453
1908	80033
1923	142830
1970	66827
2004	227934
2008	87696
2010	316991
Grand Total	1443764
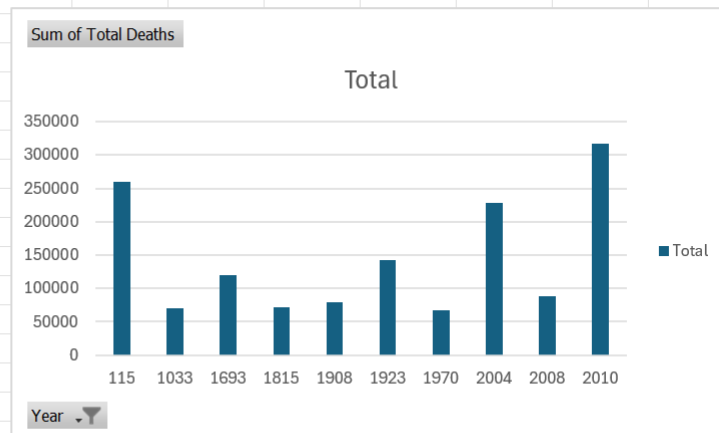
8.2.10 Top 10 of Country has Maximum Water Height

Row Labels ⊤	Sum of Maximum Water Height (m)
CHILE	391.2
CHINA	217.95
GREECE	242.93
INDONESIA	1217.67
ITALY	320.75
JAPAN	877.66
NORWAY	265.4
PERU	177.32
RUSSIA	377.76
USA	1988.37
Grand Total	6077.01
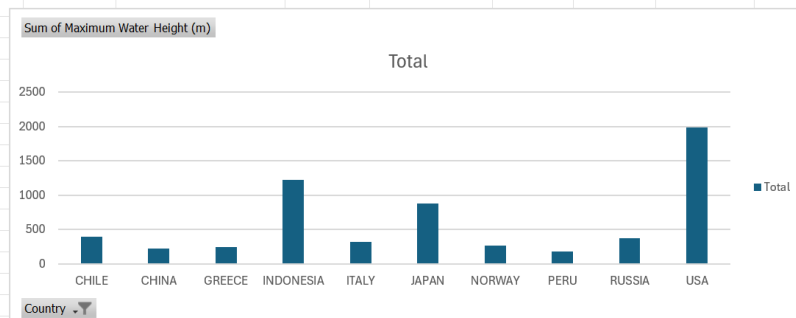
9 Data Mining

9.1 Load Dataset

```
In [2]:  # Load the dataset
         tsunami = pd.read_csv(r'C:\Users\ungdu\Downloads\Smart-Decision\Tsunami_Cut.csv')
         tsunami.head()
```

Out[2]:

	Year	Mo	Tsunami Event Validity	Tsunami Cause Code	Earthquake Magnitude	Vol	Deposits	Country	Location Name	Latitude	...	Damage ($Mil)	Damage Description	House Destroye
0	-2000	NaN	1	1.0	NaN	NaN	0	SYRIA	SYRIAN COASTS	35.683	...	NaN	4.0	Na
1	-1610	NaN	4	6.0	NaN	1351.0	28	GREECE	THERA ISLAND (SANTORINI)	36.400	...	NaN	3.0	Na
2	-1365	NaN	1	1.0	NaN	NaN	0	SYRIA	SYRIAN COASTS	35.683	...	NaN	NaN	Na
3	-1300	NaN	2	0.0	6.0	NaN	0	TURKEY	IONIAN COASTS, TROAD	39.960	...	NaN	NaN	Na
4	-760	NaN	2	0.0	NaN	NaN	0	ISRAEL	ISRAEL AND LEBANON COASTS	NaN	...	NaN	NaN	Na

5 rows × 26 columns

9.2 Visualization of Dataset
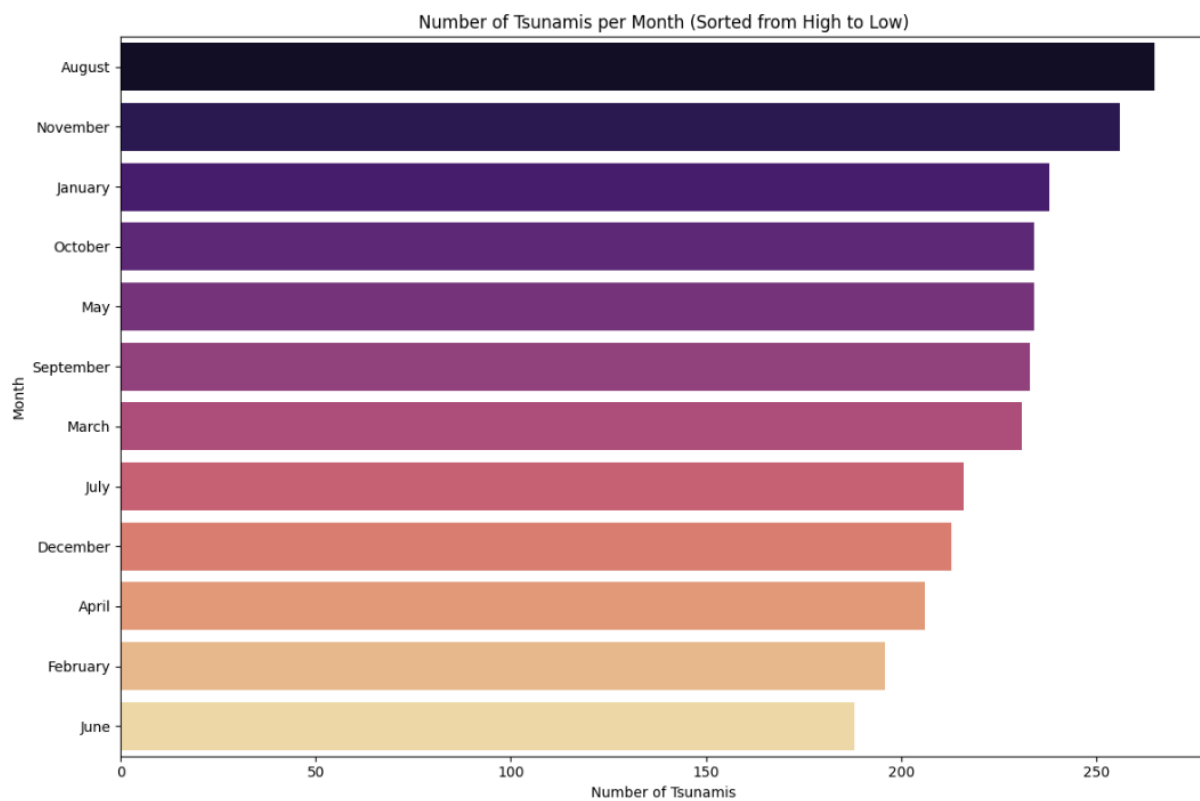
9.2.1 Visualization Number of Tsunami per Country

Number of Tsunamis per Country
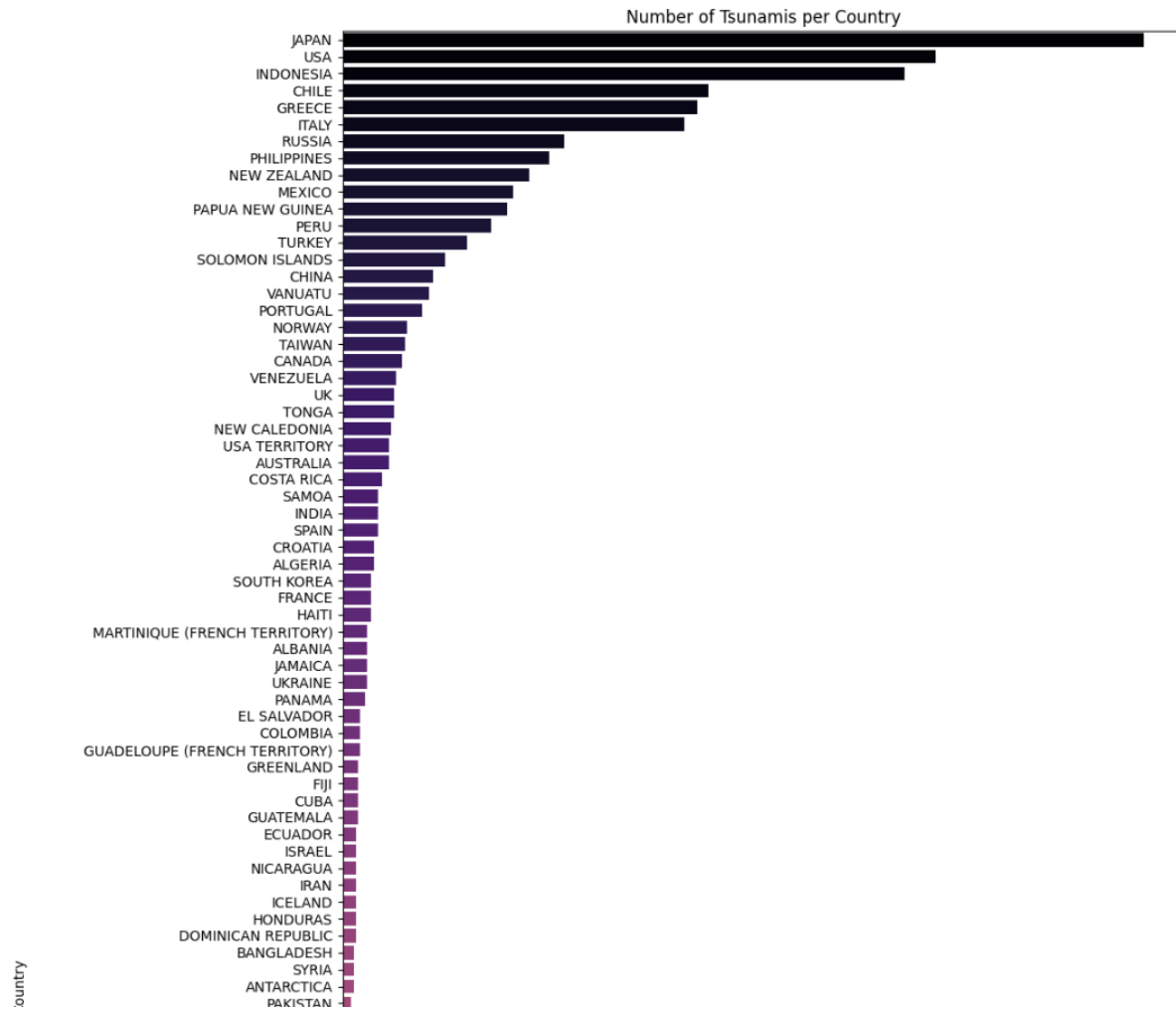
9.2.2 Visualization number of Tsunami per Cause

Number of Tsunamis per Cause

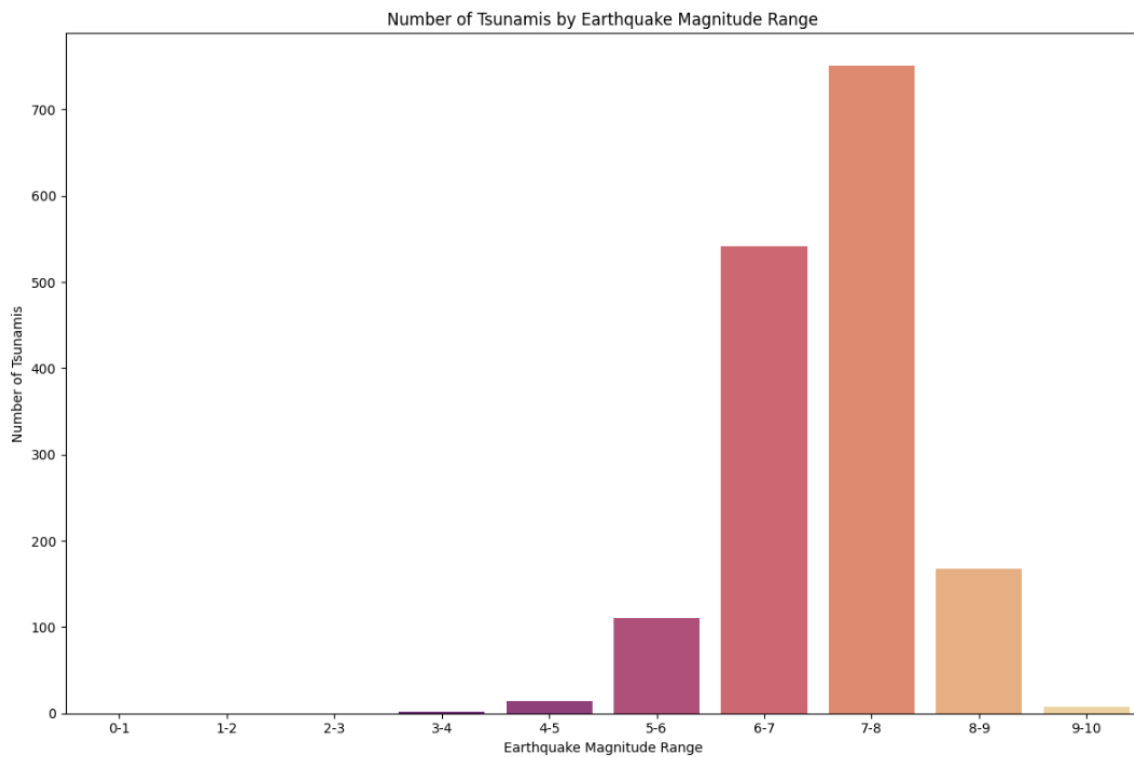### 9.2.3 Number of Tsunamis per Month (Sorted from High to Low)
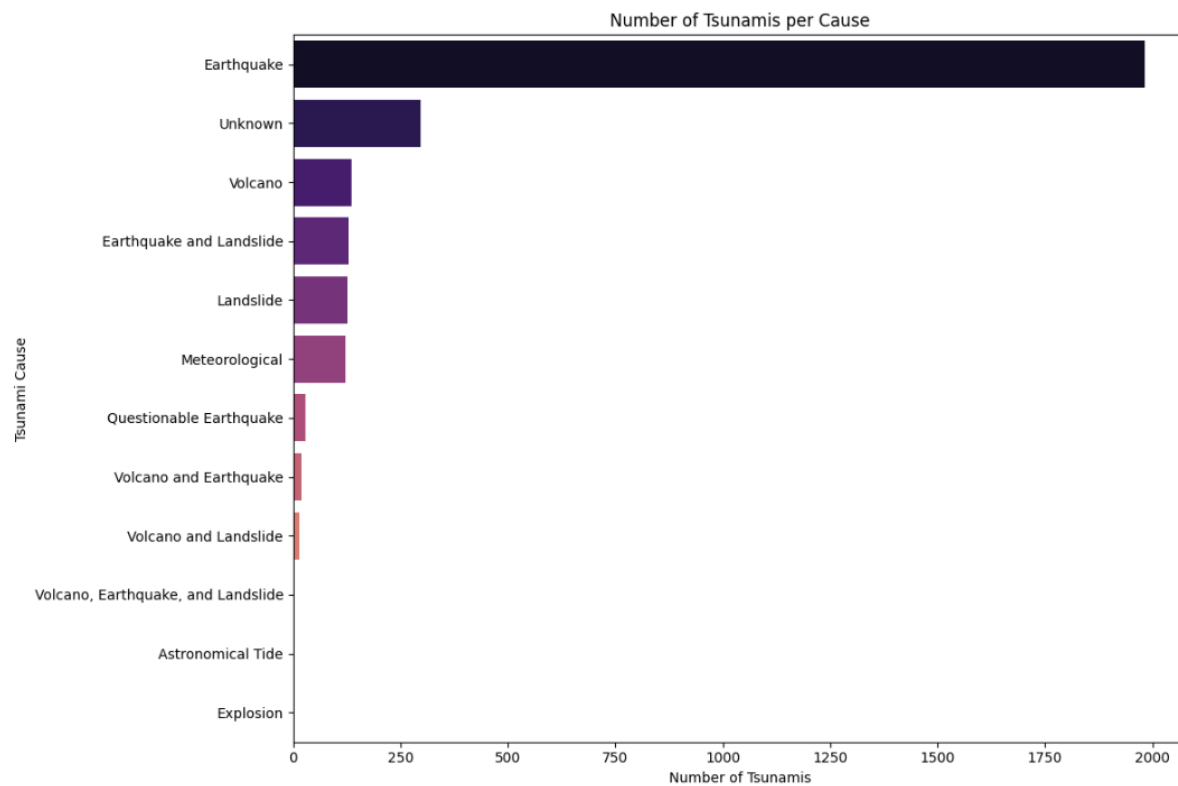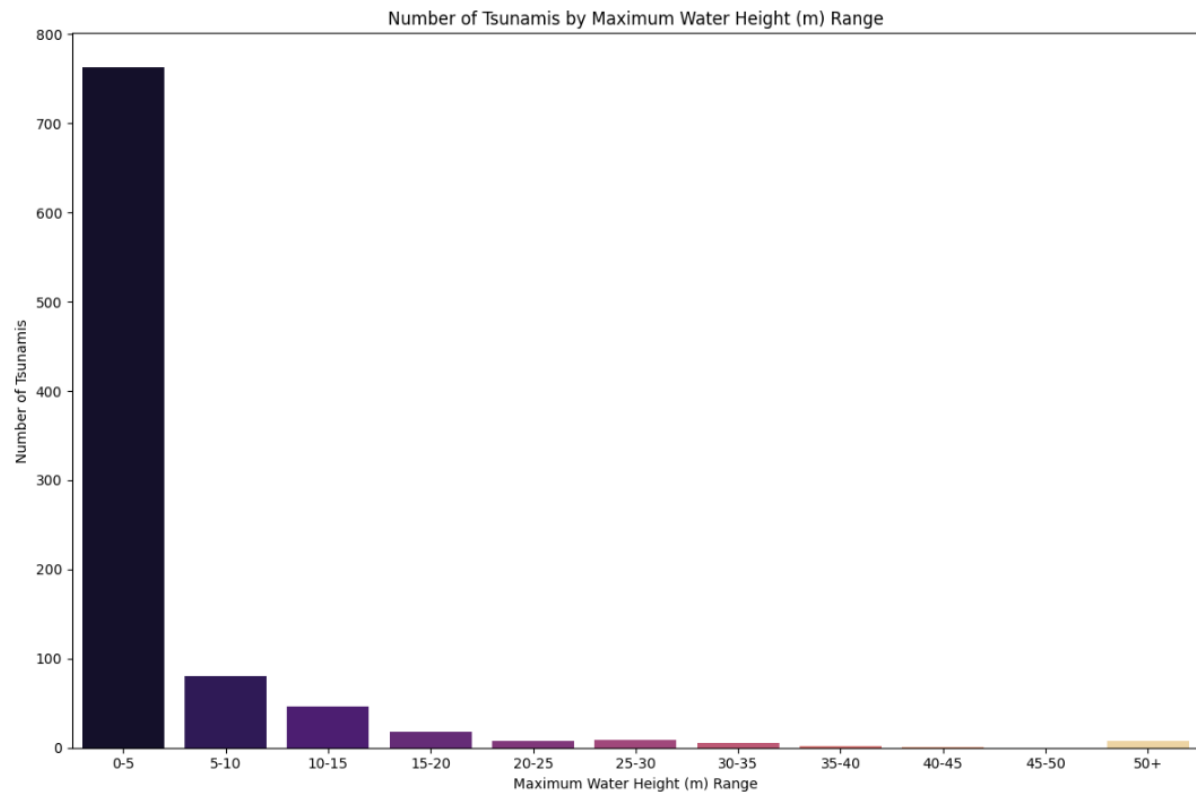

Number of Tsunamis per Month (Sorted from High to Low)

9.2.4 Number of Earthquakes by Magnitude Range


Number of Tsunamis by Earthquake Magnitude Range

9.2.5 Number of Tsunamis by Maximum Water Height (m) Range

Number of Tsunamis by Maximum Water Height (m) Range

9.2.6 Number of Tsunamis by Number of Runups Range
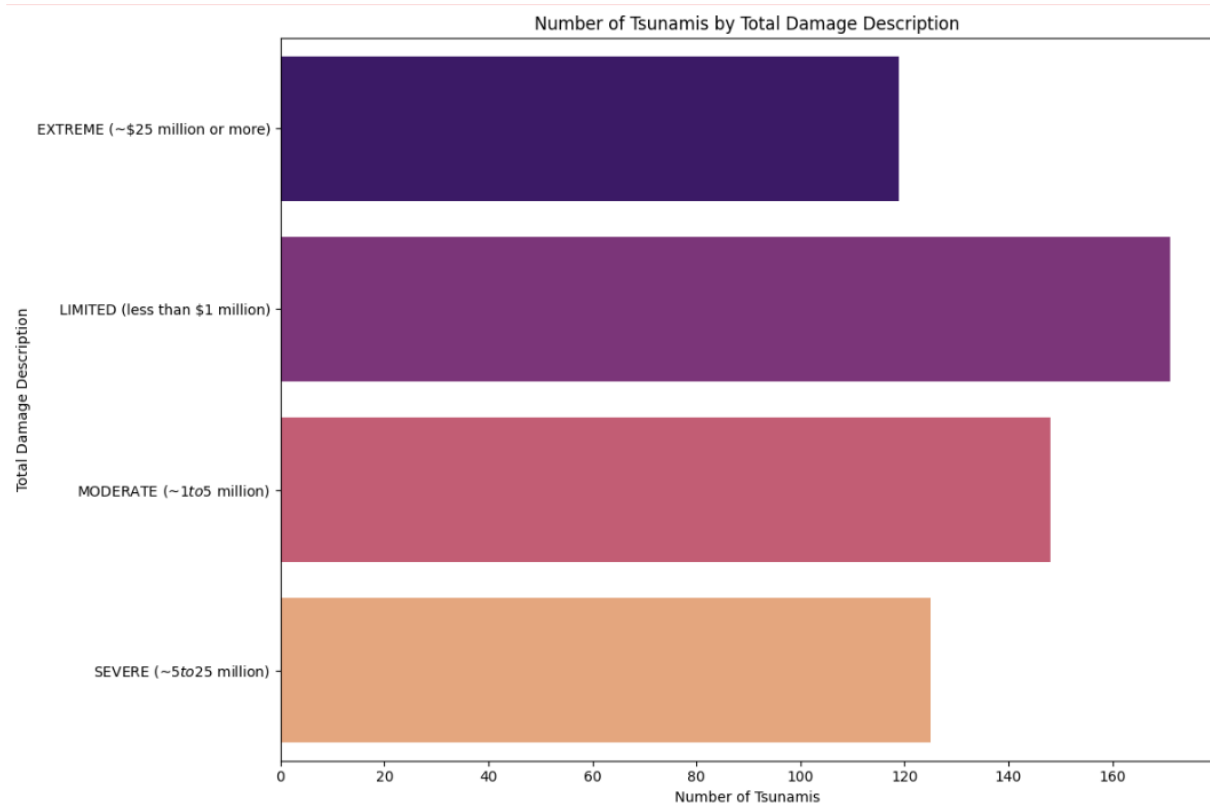

Number of Tsunamis by Number of Runups Range
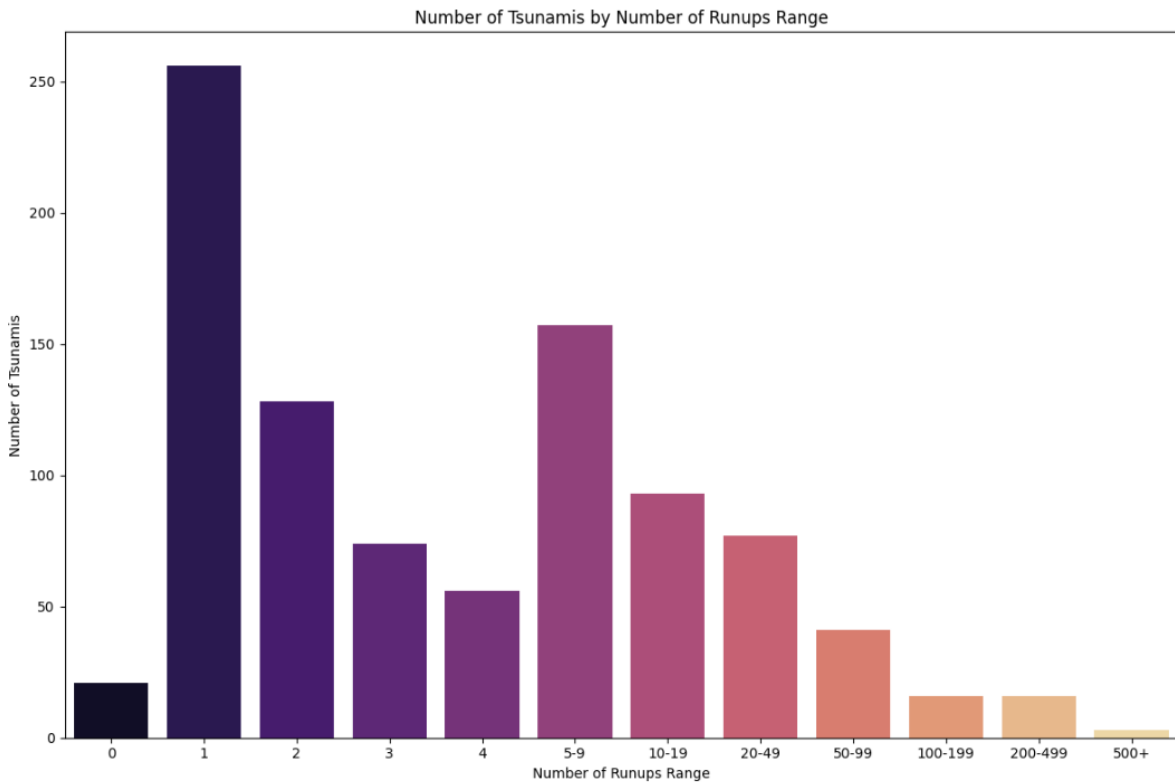
9.2.7 Number of Tsunamis by Event Validity
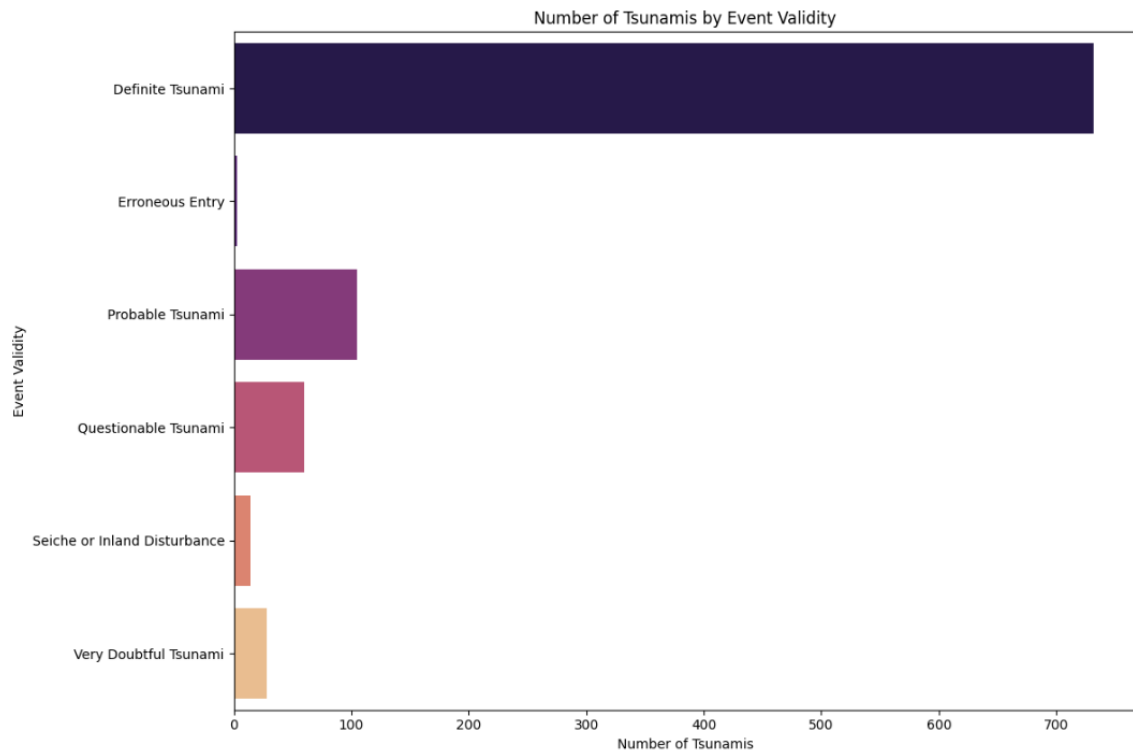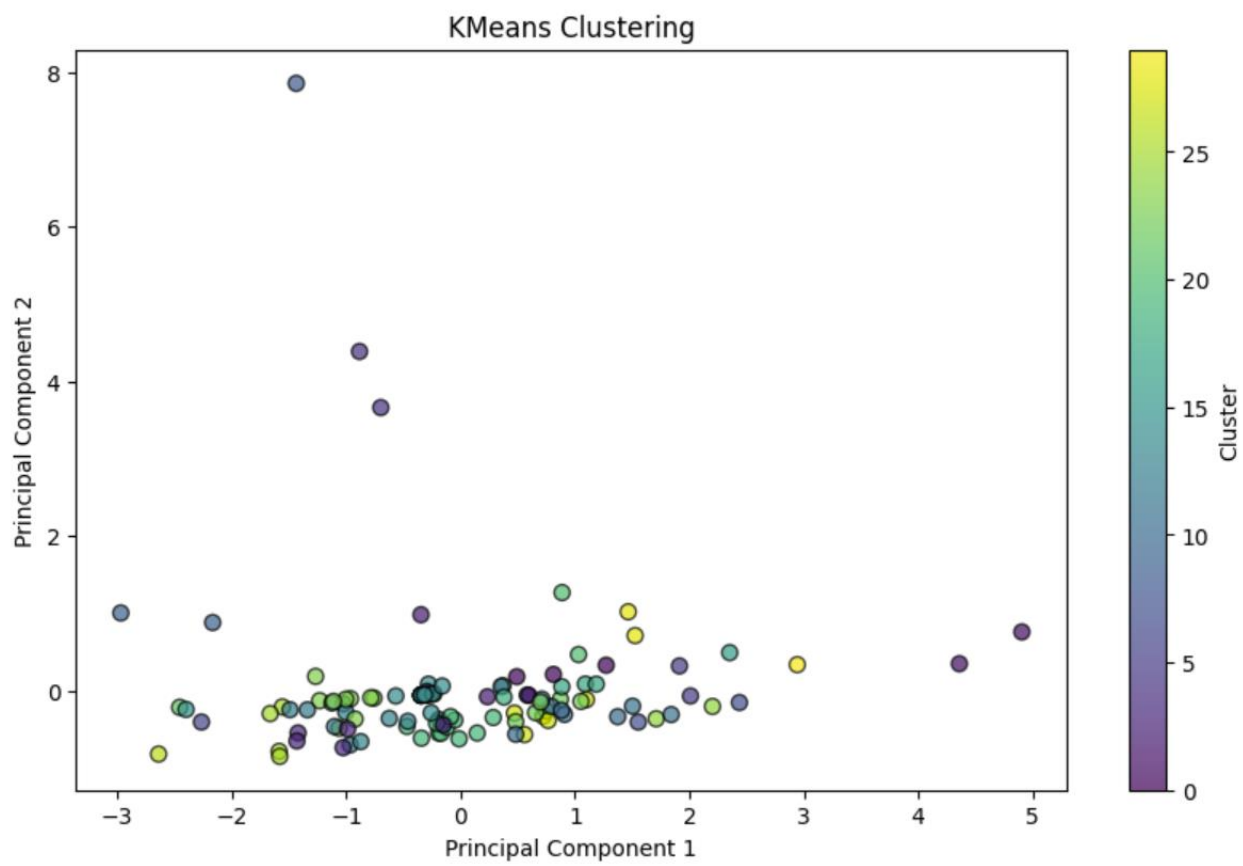

9.2.8 Number of Tsunamis by Total Death Description

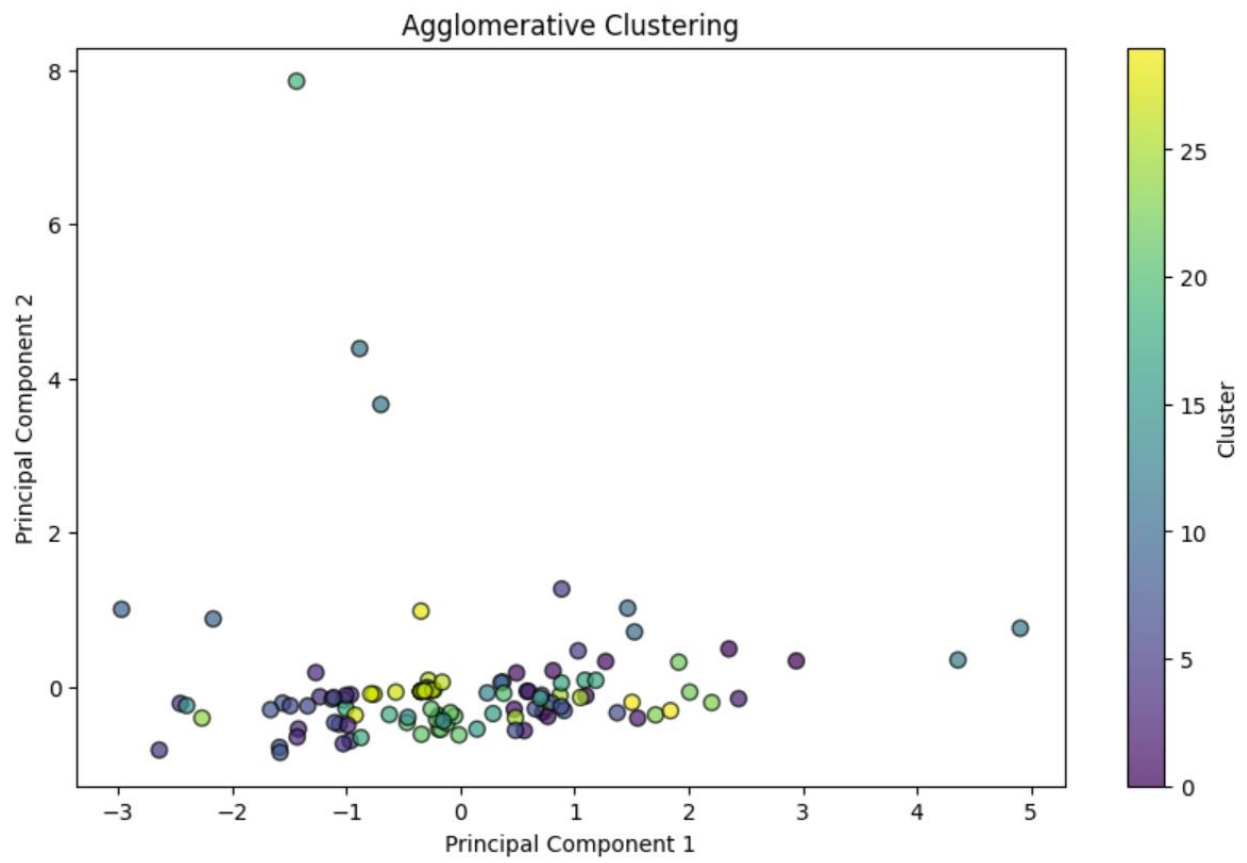### 9.2.9 Number of Tsunamis by Total Damage Description
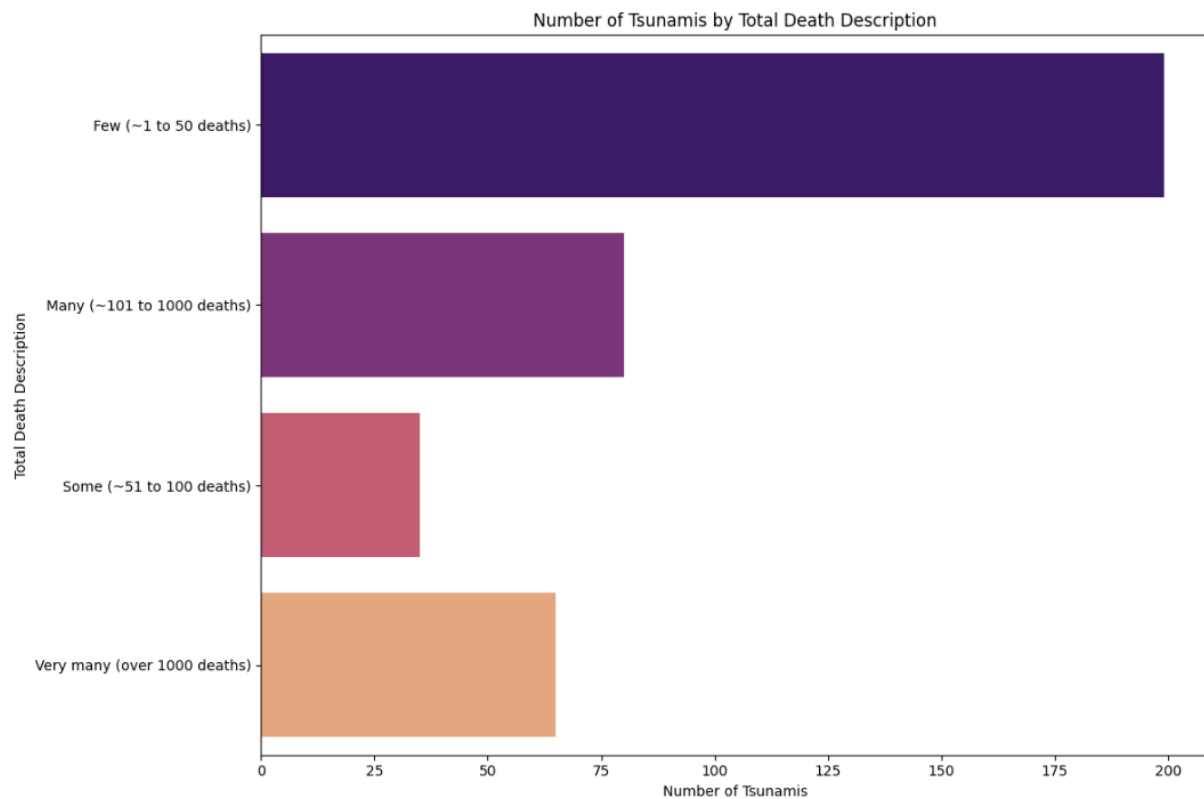
Number of Tsunamis by Total Damage Description
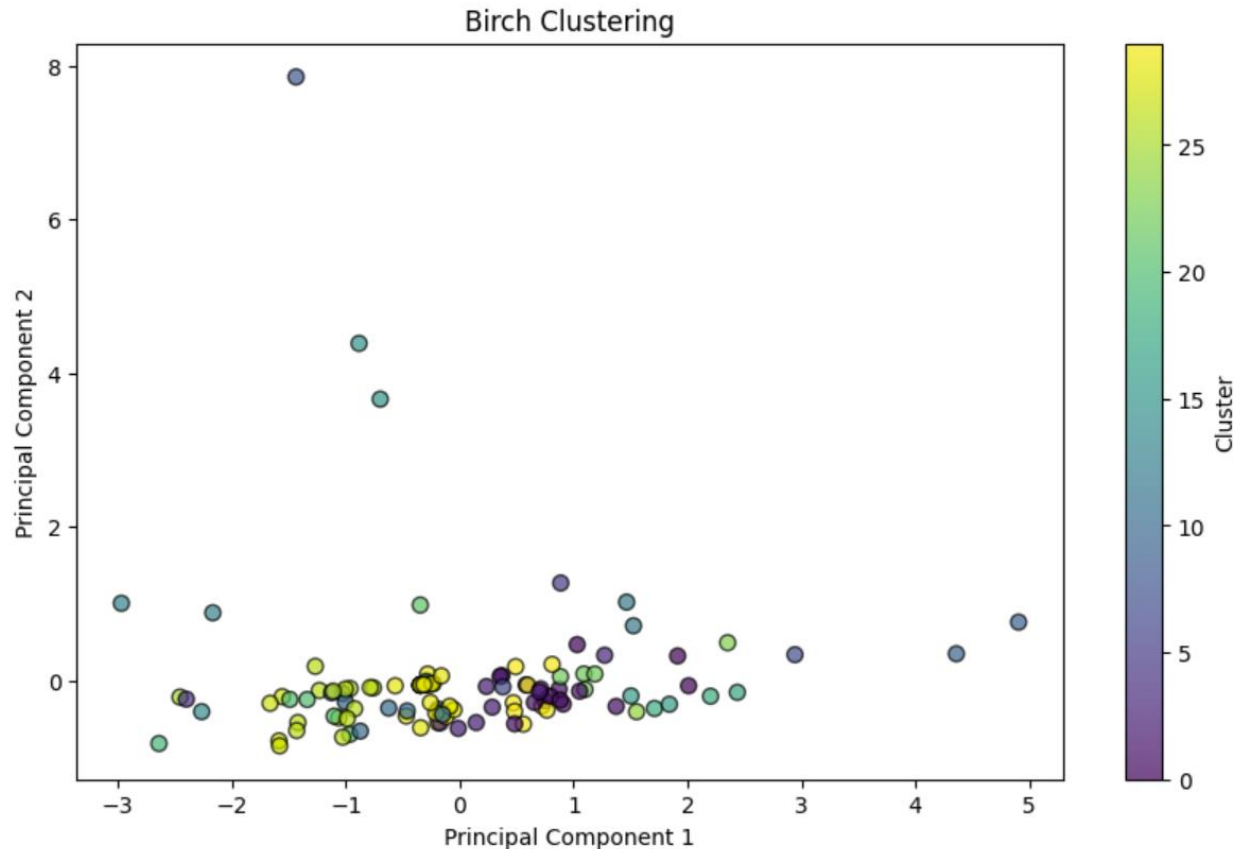
9.3 Clustering Dataset by using some algorithms
9.3.1 Clustering by using K-means

9.3.2 Clustering by using Agglomerative

Agglomerative Clustering

9.3.3 Clustering by using Birch

Birch Clustering

9.3.4 Comparison between alogirthm

Algorithm	Silhouette Score	Davies-Bouldin Index	Calinski-Harabasz Index
KMeans	0.355615	0.644754	78.74093
Agglomerative	0.389007	0.644692	84.49318
Birch	0.358131	0.5166	70.38094

Overall Conclusion:

- KMeans: Provides a good balance between cluster definition and separation but is slightly outperformed by Agglomerative.
- Agglomerative Clustering: Provides the best-defined and well-separated clusters among the three.
- Birch: Produces the most compact clusters but has less separation compared to KMeans and Agglomerative.

10 Conclusion

After analysis, we have some conclusions. Firstly, the main cause of tsunamis is earthquakes, the year with the highest number of deaths by tsunami is the 2000s. Secondly, Japan has the highest

number of Tsunami. Moreover, Most of Tsunami has Earthquake magnitude range in 7-8. Then, Tsunamis occur most frequently in August each year. Lastly, in data mining, the data used for analysis gives quite good results, highlighting visualization of data on the causes, deaths, and damages of Tsunamis. Agglomerative Clustering is the best algorithm for your data, with well-defined and well-separated clusters. KMeans also performs well but is slightly outperformed by Agglomerative. Birch provides compact clusters but with less separation between clusters.