

Challenge Hints

The following hints will assist you in completing the challenge. If you need additional help, you may view the code in the suggested answer section of the document.

Clean the Data

1. Consider using the [SCAN\(\)](#) or [SUBSTR\(\)](#) function to parse the **ID_REGIONCODE** variable.
2. Consider using the [UPCASE\(\)](#) function to standardize values of **FLAG_TSUNAMI**.
3. Consider using the [DHMS\(\)](#) function with a nested [MDY\(\)](#) function to create the **DATE_TIME** variable. Be sure to format **DATE_TIME** using the [DATETIMEw.d](#) format.
4. Consider using either the [COALESCE\(\)](#) function or the [IF-THEN/ELSE](#) statement to select the first nonmissing value in the **EQ_MAG** measurement variables.

Determine Valid and Invalid Observations

There are multiple ways to determine the valid and invalid observations. In the hint below, you determine the valid observations to write to the **Earthquakes_valid** data set, and the remaining observations are written to the **Invalid** data set. We use the [IF-THEN/ELSE](#) statement with the [AND](#) operator and with the following conditions:

- **ID:** There are multiple ways to check for a duplicate **ID** value. One way is to consider using the [FREQ](#) procedure with the [ORDER=](#) option and a [TABLES](#) statement. Find the only observation that has a frequency of 2, and use the Not Equal to comparison operator with the duplicate **ID** value to identify all valid IDs.
- **REGION_CODE:** Use the [IN](#) operator with a list of the valid region codes.
- **FLAG_TSUNAMI:** Use the [IN](#) operator with the valid values.
- **DATE_TIME:** Use the Not Equal to comparison operator to find all nonmissing date values.
- **EQ_PRIMARY:** Use [comparison](#) operators to include only the valid values.
- **FOCAL_DEPTH:** Use [comparison](#) operators to include only the valid values.

Bonus

There are multiple ways to create the **INVALID_DESCRIPTION** variable. One option is to consider using the **DATA** step with a [LENGTH](#) statement to create the character variable **INVALID_DESCRIPTION**. Then use an [IF-THEN/ELSE](#) statement to determine whether a variable is invalid. If the variable is invalid, use the [CATX\(\)](#) function to append the string that identifies the invalid variable to **INVALID_DESCRIPTION**.