1. Import Data
2. Drop all Data <1975
3. Convert Temperature Averages from ‘str’ to float
4. Replace missing Temps (in this case they were ‘-9999’ which would skew data) with np.nan
5. Divide Temps by 100 (the format for the temps was Celsius \* 100, probably used to avoid handling decimals)
6. In our data set, many countries had multiple weather stations collecting data at the same time. What we did was average all temps together across the weather stations based on year and month. Therefor, the resulting values would be one average for every month of each year 1975-2017 by country.
7. We realized a number of countries were missing continuous years of data, we decided to drop those countries with more than 1 year of missing data.
8. Then recompiled the data into a DF again by Country/Year/Month