

CSCI 1100 — Computer Science 1 Homework 7

Dictionaries

Homework Overview

This homework is worth **90 points** toward your overall homework grade and is due **Thursday November 12, 2015 at 11:59:59 pm**. It consists of one single program called `hw7.py` that assumes the presence of a data file called `tempdata.json`.

Just in time for these extremely warm November days, we have a climate based homework. The data for this homework comes from NOAA (National Oceanic and Atmospheric Administration). It is temperature data from Troy, NY from 1956 till June of 2015, measured at Hudson River Lock and Dam (you can see it for yourself here: <http://www.google.com/maps/place/42.75,-73.68333>). The data is given to you in a file called `tempdata.json` in JSON format, which can be read in its entirety with the following three lines of code:

```
import json

if __name__ == '__main__':
    data = json.loads(open("tempdata.json").read())
```

The data variable read this way is a list of dictionaries. Each dictionary in the list `data` contains climate values for a single month of the year. For example:

```
>>> print data[0]
{'u'MMNT': -9999, u'TPCP': -9999, u'MMXT': -9999, u'DX32': 0, u'MNTM': 0, u'DT90': 0,
 u'EMNT': 51, u'MXSD': -9999, u'DT32': 0, u'month': 7, u'DP05': -9999, u'year': 1956,
 u'DSNW': -9999, u'DP01': -9999, u'TSNW': -9999, u'DP10': -9999, u'DT00': 0, u'EMXT':
 -9999}
```

Each attribute of the dictionary is a climate measurement for a given year and month. The `u` before each attribute name means that these are strings encoded in Unicode, for example `u'MMNT'`. You can treat these strings as any regular string.

Here is the list of attributes each dictionary has and their meaning:

month:	Month of year (1-12)
year:	Year
EMNT:	Extreme minimum daily temperature
EMXT:	Extreme maximum daily temperature
MMNT:	Monthly Mean minimum temperature
MMXT:	Monthly Mean maximum temperature
MNTM:	Monthly mean temperature
MXSD:	Maximum snow depth
TPCP:	Total precipitation
TSNW:	Total snow fall
DT00:	Number days with minimum temperature less than or equal to 0.0 F
DT32:	Number days with minimum temperature less than or equal to 32.0 F
DT90:	Number days with maximum temperature greater than or equal 90.0 F
DX32:	Number days with maximum temperature less than or equal to 32.0 F
DP01:	Number of days with greater than or equal to 0.1 inch of precipitation
DP05:	Number of days with greater than or equal to 0.5 inch of precipitation
DP10:	Number of days with greater than or equal to 1.0 inch of precipitation

Note that some values are missing, especially for older data. In this case, you have a special value (often -9999). You are expected to disregard any value that is zero or -9999.

Problem Description

Write a program that reads a certain month from the user (an index between 1 and 12), and find the data for this month in each year to print out the following information:

- Top 3 years and associated values for:
 - highest maximum temperature (EMXT)
 - lowest minimum temperature (EMNT)
 - highest number of days with maximum temperature above 90 (DT90)
 - highest number of days with maximum temperature below 32 (DX32)
 - highest total precipitation (TPCP)
 - lowest total precipitation (TPCP)
 - highest snow depth (TSNW)
 - lowest snow depth (TSNW)

If there are less than 3 values above zero for a given parameter for that month, then print that there is not enough data (for example snow depth in August!).

- Average of mean temperature across different years (MNTM)
 - Overall average (average of all years for MNTM)
 - Average for the first 10 years we have data for (for MNTM)
 - Average for the last 10 years we have data for (for MNTM)
- A histogram of (integer part of) overall average temperatures in 10 year increments (plus the last remaining years after counting by 10s), print a single star for each degree. Note that the histogram should disregard missing values.

Hint. Just a few hints on how to proceed in solving this homework. It sounds like there are many things to compute, but it is actually the same thing for many different keys of the same dictionary. My program is 100 lines, it basically has two main functions.

The first function loops through through the dictionary for any given month and key (both parameters to my function) and computes all the above statistics, and returns them in a dictionary. Then, in my program, I simply need to print these values. As all the above computation really uses the same set of value and year pairs, this saves a lot of repetition. In this dictionary, I even return the list of values.

I have a second function for printing the histogram of a set of values. I can simply use the dictionary returned by the first function to find the values, and then use them to print the histogram.

This is a good homework to fully make use of generic functions and dictionaries to help you solve many different problems. Start with one function to print the first line of the output (for EMXT), then think about how to generalize it to solve all the other parts quickly.

Expected output

Below you can see the expected functioning of this program with the file we gave you (note: we might change the file in the homework submission server):

```
Enter a month (1-12) => 11
11
Temperatures
-----
Highest max  value => 1975: 83.0, 1982: 81.0, 1956: 80.0
Lowest min value => 1958: 10.0, 1996: 10.0, 1957: 12.0
Highest days with max >= 90 => Not enough data
Highest days with max <= 32 => 1989: 4.0, 2013: 3.0, 2005: 3.0

Precipitation
-----
Highest total => 1972: 6.9, 1969: 6.8, 1959: 6.3
Lowest total => 2011: 0.4, 1978: 0.7, 2012: 0.9
Highest snow depth => 1971: 22.5, 1972: 19.0, 1985: 11.5
Lowest snow depth => 2000: 0.1, 1962: 0.3, 2013: 0.5

Average temperatures
-----
Overall: 41.0
First 10 years: 41.4
Last 10 years: 41.2

1956-1965: *****
1966-1975: *****
1976-1985: *****
1986-1995: *****
1996-2005: *****
2006-2014: *****
```

Enter a month (1-12) => 7

7

Temperatures

Highest max value => 1995: 101.0, 2010: 99.0, 1988: 99.0
Lowest min value => 1963: 45.0, 1960: 46.0, 1961: 46.0
Highest days with max >= 90 => 1988: 14.0, 2010: 12.0, 1983: 12.0
Highest days with max <= 32 => Not enough data

Precipitation

Highest total => 1984: 9.4, 2004: 8.7, 2009: 8.4
Lowest total => 1964: 1.1, 2002: 1.4, 1977: 1.9
Highest snow depth => Not enough data
Lowest snow depth => Not enough data

Average temperatures

Overall: 73.0
First 10 years: 72.0
Last 10 years: 73.9

1957-1966: *****
1967-1976: *****
1977-1986: *****
1987-1996: *****
1997-2007: *****
2008-2014: *****
