**Design of Rainfall Statistics System**

Author: Vivek Makarabooshanam

Date: 09/27/2015

Version: 1.0

**Problem:**

Write a Java program to compute sorted MSA by wettest population.

Original problem statement: <https://docs.google.com/document/d/1nosltNbyQCy-TDKo15xa4aboZpbsOTm3QMDXLfGWPKc/edit?usp=sharing>

**Assumptions:**

1. Since Statistical Area – Population file does not have current year (2015) population, I projected the population for 2015 by using 2010 population as base and by averaging the growth percentage.
2. Since MSA did not tie well with WBAN IDs just by location, I downloaded one more dataset (List1.xls) to map the WBAN to County to MSA <http://www.census.gov/population/metro/data/def.html>. This file is present in here: src/main/resources/MSA-County-Census-List1.xls

**Input Data:**

I chose the following files and corresponding data points to build the results:

|  |  |  |  |
| --- | --- | --- | --- |
| File Name | CPH-T-5.xls | | |
| Description | Statistical Area – Population for 2000 & 2010 | | |
| Data Points used | Title (Statistical Area Name) | April 1, 2010 Population | Percent |

|  |  |  |  |
| --- | --- | --- | --- |
| File Name | MSA-County-Census-List1.xls | | |
| Description | Statistical Area and corresponding County | | |
| Data Points used | CBSA Title (Statistical Area Name) | County/County Equivalent | State |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| File Name | wbanmasterlist.psv | | | |
| Description | WBAN Master list | | | |
| Data Points used | WBAN\_ID | Station Name | County | State\_Province |

|  |  |  |  |
| --- | --- | --- | --- |
| File Name | 201505station.txt | | |
| Description | List of WBAN stations measured in May 2015 | | |
| Data Points used | WBAN (ID) | Name | State |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| File Name | 2015precip.txt | | | | |
| Description | Precipitation data – format <http://www.ncdc.noaa.gov/qclcd/qclcdhrlyprecip.htm> | | | | |
| Data Points used | WBAN (ID) | YearMonthDay  (yyyyMMdd) | Hour (0 – 24 hrs format) | Precipitation (inch) or T for trace | PrecipitationFlag |

**High level Design:**

Component Diagram and Usage of Pipe-Filter pattern built into Java 8.

Various high level entities in system: Entity, DataLoader & Calculator & Filters

**Class Diagram:**

Census related entites: USAState, USACity, USACounty, USAStatisticalArea

Weather related entities: WBAN

**Sequence Diagram:**

DataLoader.load() – Calculator.calculate()

**Programmer Testing:**

Junit4 base testsuite:

TestPopulationWetnessCalculator.testAustinAreaRainfallWithGoodAndBadWBANs()

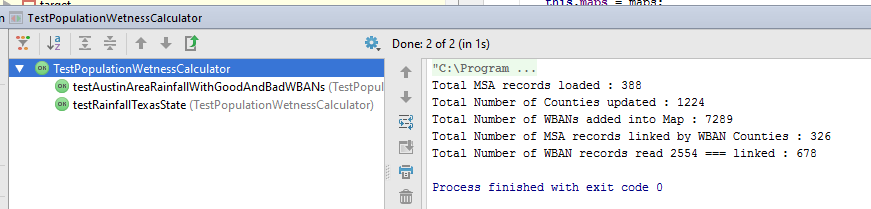
TestPopulationWetnessCalculator.testRainfallTexasState

Both testcases use below mock test dataset for Austin-RoundRock area along with invalid WBANs:

****

Test Execution Output:

****



**Execution Output:**

Here is the output for both statistic using actual Rainfall dataset:

Total MSA records loaded : 388

Total Number of Counties updated : 1224

Total Number of WBANs added into Map : 7289

Total Number of MSA records linked by WBAN Counties : 326

Total Number of WBAN records read 2554 === linked : 678

Context Data Load Time :1301 ms

PopulationWetnessByMSA Calculation Time :918 ms

RainfallByState Calculation Time :1679 ms

**PopulationWetnessByMSA – Statistic output file:**



**RainfallByState – Statistic output file:**

****

**Ref:**

<http://www.ncdc.noaa.gov/qclcd/qclcddocumentation.pdf>