DATA-51100: Statistical Programming
Programming Assignment 1 – Online Descriptive Statistics

Introduction
Descriptive statistics are quantities used for summarizing a data sample. Examples include the *sample mean* and the *sample variance*. Given a set of samples, $x_1, x_2, ..., x_n$, the sample mean is defined as

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \ .$$

The (unbiased) sample variance is then defined as

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \ .$$

In certain applications (e.g. real-time systems), it is important to be able to compute these quantities using an *online algorithm*. An online algorithm computes as the data is coming in, piece-by-piece, rather than waiting until all the data is available. The advantage is that it doesn't have to store all the data up to the current time to compute the mean and variance. Rather, it can simply update those quantities based on the current most entered value.

An online algorithm for computing the sample mean and variance starts with $\bar{x}_n = 0$ and $s_n^2 = 0$, where $\bar{x}_n$, s_n^2 are the sample mean and variance for $n = 0$ points, respectfully. Then, for each new sample point, the quantities are updated using the following formulas:

$$\bar{x}_n = \bar{x}_{n-1} + \frac{x_n - \bar{x}_{n-1}}{n}$$

and

$$s_n^2 = \frac{n-2}{n-1} s_{n-1}^2 + \frac{(x_n - \bar{x}_{n-1})^2}{n} \quad for \ (n > 1).$$

were,

n is the number of values entered by user so far

$\bar{x}_{n-1}$ is the current mean after the $n-1$'th value was entered

$\bar{x}_n$ is the updated mean after the n'th value is entered

s_{n-1}^2 is the variance after the $n-1$'th value was entered

s_n^2 is updated variance after the n'th value is entered

Requirements
You are to create a program using Python that asks the user for a nonnegative number, then computes the mean and variance using the above given online update formulas which should be displayed on the console screen. The program should end when a user enters a negative number.

Additional Requirements

1. The name of your source code file should be `OnlineStats.py`. All your code should be within a single file.
2. Your code should follow good coding practices, including good use of whitespace and use of both inline and block comments.
3. You need to use meaningful identifier names that conform to standard naming conventions.
4. At the top of each file, you need to put in a block comment with the following information: your name, date, course name, semester, and assignment name.
5. The output of your program should **exactly** match the sample program output given at the end. That is, for same input, it should generate the same output. Note that I may use other test cases for grading your program and your code needs to work correctly in all cases.
6. The program needs to compute the mean and variance using the online algorithm. That is, you cannot store all the entered values in a list. Instead, just update the mean and variance for each new entered value using the update formulas.

What to Turn In

You will turn in a screenshot of your output and the single OnlineStats.py file using BlackBoard.

```
CPSC-51100, [semester] [year]
NAME: [put your name here]
PROGRAMMING ASSIGNMENT #1

Enter a number: 1
Mean is  1.0  variance is  0

Enter a number: 2
Mean is  1.5  variance is  0.5

Enter a number: 3
Mean is  2.0  variance is  1.0

Enter a number: 4
Mean is  2.5  variance is  1.66666666667

Enter a number: 5
Mean is  3.0  variance is  2.5

Enter a number: 6
Mean is  3.5  variance is  3.5

Enter a number: 7
Mean is  4.0  variance is  4.66666666667

Enter a number: 8
Mean is  4.5  variance is  6.0

Enter a number: 9
Mean is  5.0  variance is  7.5

Enter a number: 10
Mean is  5.5  variance is  9.16666666667

Enter a number: -1
```

```
CPSC-51100, [semester] [year]
NAME: [put your name here]
PROGRAMMING ASSIGNMENT #1

Enter a number: 8
Mean is  8.0  variance is  0

Enter a number: 7
Mean is  7.5  variance is  0.5

Enter a number: 12
Mean is  9.0  variance is  7.0

Enter a number: 0
Mean is  6.75  variance is  24.9166666667

Enter a number: -1
```