Xi Zhao (NUID: 001054333)

INFO 6205

Program Structures & Algorithms

Fall 2020

Section 6

Assignment 3

- **Task:** Using your implementation of UF_HWQUPC, develop a UF ("union-find") client that takes an integer value n from the command line to determine the number of "sites." Then generates random pairs of integers between 0 and n-1, calling connected() to determine if they are connected and union() if not. Loop until all sites are connected then print the number of connections generated. Package your program as a static method count() that takes n as the argument and returns the number of connections; and a main() that takes n from the command line, calls count() and prints the returned value. If you prefer, you can create a main program that doesn't require any input and runs the experiment for a fixed set of n values. Show evidence of your run(s).

- **Output (Over 10 times run of each N value, 90 N values in total)**
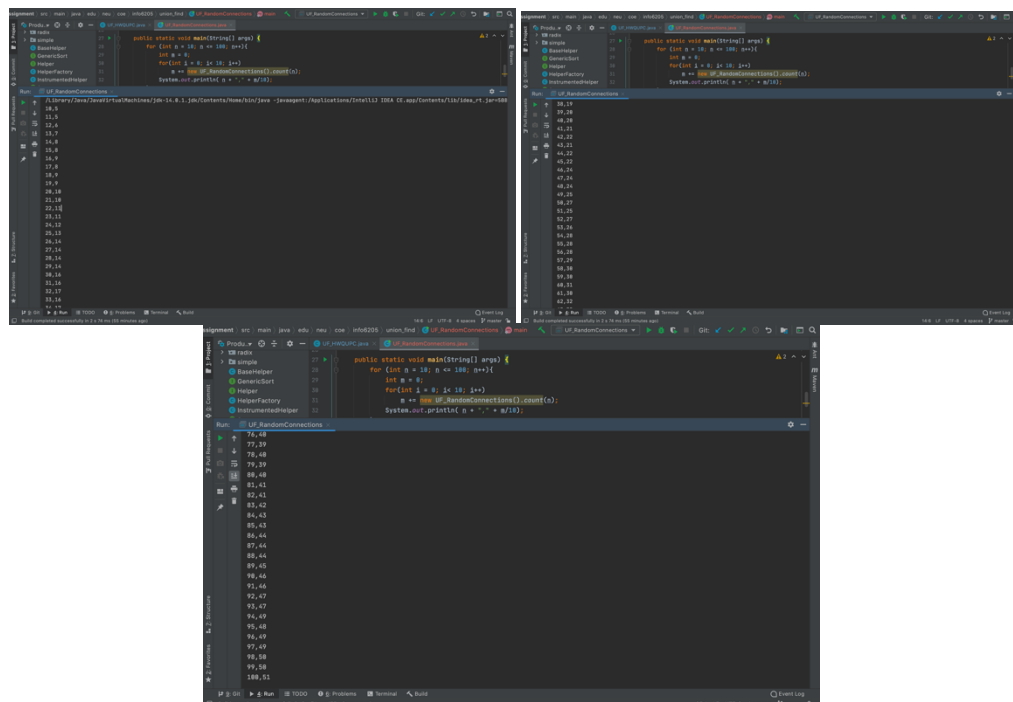
- **Relationship conclusion**

  The relationship between the given N value and connections is linear. Roughly, the formula of the number n(x) and the connection number(y) is
  y =0.5087x - 0.0205**.**

- **Evidence to support relationship**

  By using Excel to analyze and deduce the relationship of the two numbers, scatter plot with different trend lines are generated to prove the relationship of the two numbers. At the same time, the correlation test which is used to evaluate the association between two or more variables is automatically conducted along with the trend line. Correlation coefficient R is comprised between -1 and 1:

    i.   -1 indicates a strong negative correlation: this means that every time x increases, y decreases;
    ii. 0 means that there is no association between the two variables (x and y);
    iii. 1 indicates a strong positive correlation: this means that y increases with x.
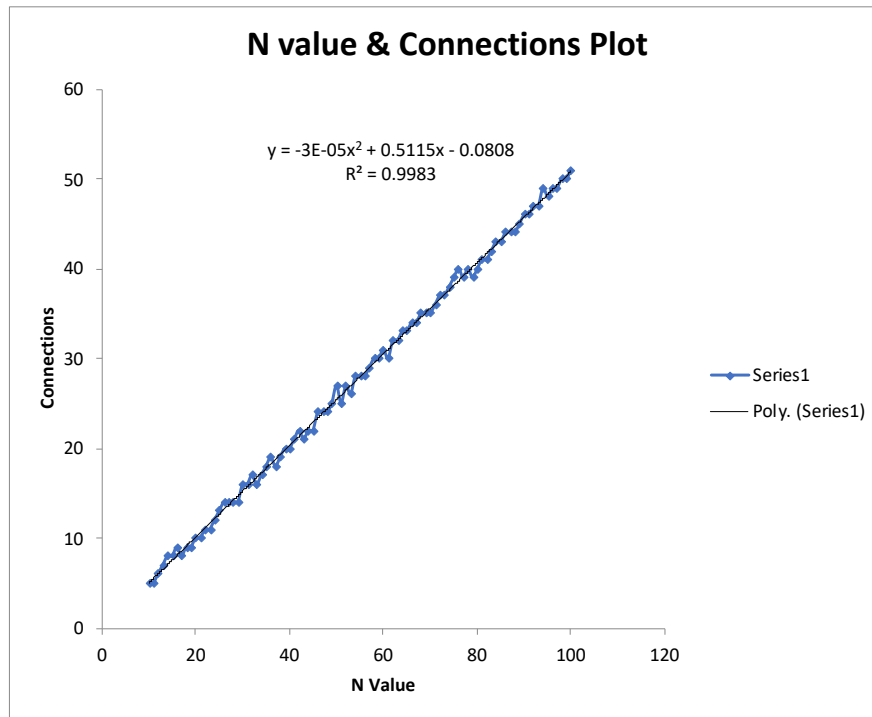
  Usually, R^2 is used to demonstrate the correlation degree. R^2 is bigger and the correlation is stronger.

  In this case, there are 4 highly possible trend lines which illustrate different possible correlation between the two numbers:
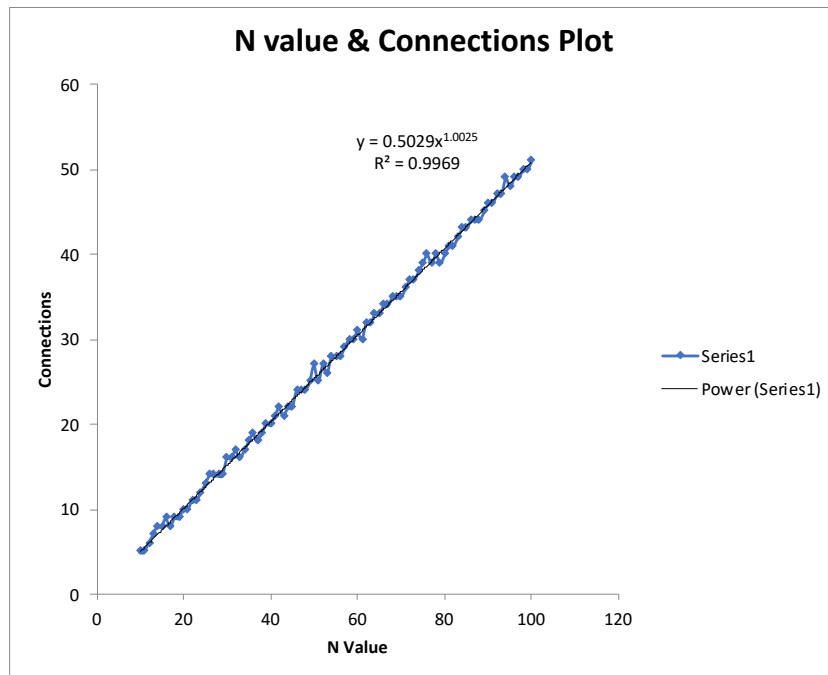
    1. Linear, y = 0.5087x - 0.0205
        R^2 = 0.9983
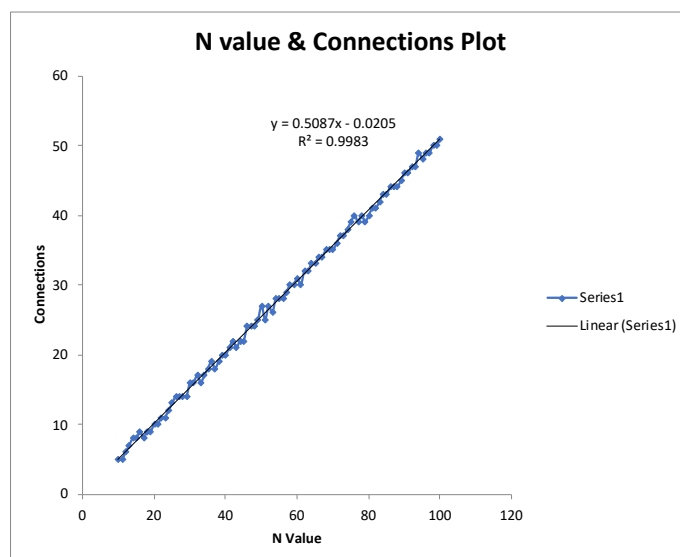
2. Polynomial, $y = -3E-05x2 + 0.5115x - 0.0808$
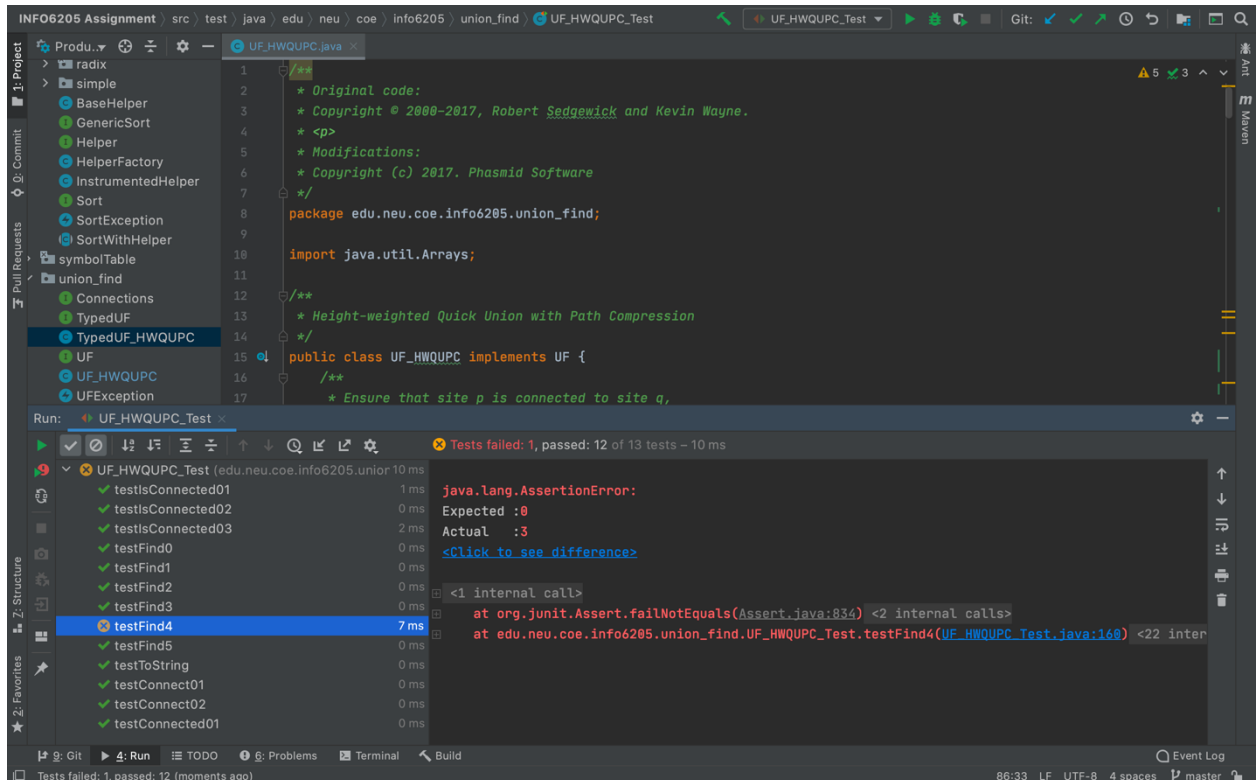 $R^2 = 0.9983$

3. Power, $y = 0.5029x1.0025$
   $R^2 = 0.9969$

Xi Zhao (NUID: 001054333)

The highest R^2 is the linear trend line. Therefore, the highest possible correlation of the n number and the connections UF generates is linear correlation.

- **Screenshot of Unit test passing**

- **Reference:**

Taylor, R. (1990). Interpretation of the Correlation Coefficient: A Basic Review. *Journal of Diagnostic Medical Sonography, 6*(1), 35-39. doi:10.1177/875647939000600106