DP2 2020-2021

Performance Report

June Group #4

Repository

<https://github.com/rodddella/Acme-Planner.git>

Members

* Guillermo Diz Gil
* Francisco Rodríguez Pérez
* Carmen María Muñoz Pérez
* George Laurentiu Bogdan

Tutor

* Rafael Corchuelo Gil

Contenido

[Summary 3](#_Toc73585850)

[Computer’s configuration 3](#_Toc73585851)

[Computer 1: Guillermo’s computer 3](#_Toc73585852)

[Computer 2: George’s computer 3](#_Toc73585853)

[Obtaining the data 4](#_Toc73585854)

[Analyzing the data 5](#_Toc73585855)

[Hypothesis contrast 6](#_Toc73585856)

# Summary

This document details the procedure followed to analyze the performance of the project on two different computers, from the collection of the data to the result of the analysis.

# Computer’s configuration

## Computer 1: Guillermo’s computer

The computer labeled as PC1 in the analysis corresponds to Guillermo’s computer. It has the following specs:

* CPU: Intel Core i5 9400F
* RAM: 16GB
* Operating system: Windows 10

## Computer 2: George’s computer

The computer labeled as PC2 in the analysis corresponds to George’s computer. It has the following specs:

* CPU: Intel Core i7 4790k
* RAM: 16GB
* Operating system: Windows 10

# Obtaining the data

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente con confianza mediaTo obtain the performance data when responding to requests, all the tests developed in the previous sprint were run on each of the test computers. In this way, a sample file in CSV format has been generated for each of them, containing a time stamp, the path and the time it took the server to provide a response. For each file, a total of 1783 samples were generated.

Texto

Descripción generada automáticamente

Since the CSV format is not convenient for data analysis, both files have been imported into Excel format, sorting them in a way that makes them easier to consult.

Interfaz de usuario gráfica, Aplicación, Tabla, Excel

Descripción generada automáticamente

# Analyzing the data

Once the data were obtained in a convenient format, the Excel data analysis wizard was used to obtain a statistical summary of each of the sample files using the wall times obtained.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Among other data, we obtain the mean response times, the variance, and the width of the 95% confidence interval, which will later be used to confirm whether our data exceed the defined maximum response time threshold. The confidence interval will be calculated as an interval whose left end is the mean of the data minus the width of the confidence interval, and whose right end is the mean of the data plus the obtained width.

Interfaz de usuario gráfica, Aplicación, Tabla, Excel

Descripción generada automáticamente

As can be seen, a threshold of 500 milliseconds maximum response time has been defined. Since both computers have managed to stay below the threshold, there is no need to refactor the developed code.

# Hypothesis contrast

To perform the hypothesis testing, we will load the data obtained previously in a new Excel sheet together with its statistical summary.

Interfaz de usuario gráfica, Aplicación, Tabla

Descripción generada automáticamente

There are several ways to perform the hypothesis testing, and we must choose one taking into account the number of samples we have obtained. Since we have obtained 1783 for each computer that has run the tests, we will perform the Z test, which is the option to follow when we have more than 50 samples.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

Since we have run the tests on two computers with different processing power, we expect the Z-test to determine that the samples obtained are different. For this, we will consider as a key value the P-value, which we will compare with the statistical significance (alpha), calculated as one minus the confidence percentage chosen above (0.05 in our case). If the P-value is lower, we will have demonstrated that the sample means are different.

Tabla

Descripción generada automáticamente

Thus, we can see how the P-value has turned out to be 0, which is lower than the expected alpha value. This verifies the hypothesis that the means are different since the execution on both computers was not identical. No further steps are necessary.