**Introduction**[**¶**](https://render.githubusercontent.com/view/ipynb?commit=66b9197211dc8493d6435f4e84a15bd627677327&enc_url=68747470733a2f2f7261772e67697468756275736572636f6e74656e742e636f6d2f43617374646561746839372f74657261706978656c2d6576616c756174696f6e2f363662393139373231316463383439336436343335663465383461313562643632373637373332372f6e6f7465626f6f6b732f7772697474656e2d7265706f72742d616d6d61722d3135303435343338382e6970796e62&nwo=Castdeath97%2Fterapixel-evaluation&path=notebooks%2Fwritten-report-ammar-150454388.ipynb&repository_id=163946192&repository_type=Repository#Introduction)

The report analyses the terapixel processing GPUs as a part of the cloud computing project. The report covers the project backgrounds, objectives and findings. It also summarises the processes and tools used in this project. CRISP DM approach is used in the project.

**Project Background and Justification**

Terapixel and high-resolution imaging reconstruction and tiling is becoming an important tool in many fields. It is helpful in creating high resolution 3D image of real time images. Since the terapixel processing requires high resolution it needs high performance infrastructures like high end GPUs. These high-end infrastructures are generally offered by the cloud services which allows scalability and optimization. Hence improving and optimizing the usage of the GPU performance is essential to improve the terapixel complex processing.

.

**Project Criteria/Objectives**

The objective is to find the trends of the usage of GPUs which is deployed for terapixel rendering on azure platform. Here the usage of memory, how the workloads are distributed and the how the GPUs perform to heavy loading and their activity is analysed. We analyse the following below:

* Analyse how well the cloud application parallelises the workload
* Measure and analyse memory usage
* Measure how well the GPUs respond to varying data scales and loads

**Tools, Methodology and Evaluation Design**

The project will follow the CRISP-DM process model. Git and Github is used for version control to manage the progress of project template.

The project will be divided into these stages:

1. Business Understanding:

Identifying the purpose of the data and determining the best method to approach the problem.

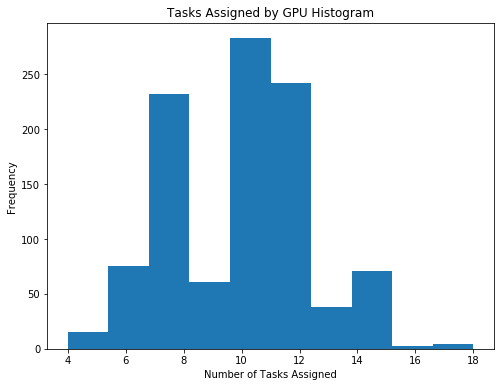
1. Data Preparation:
   * Cleaning the data to be used to build the model with necessary changes.
2. Data Understanding:
   * Performing the basic EDA of the dataset provided after cleaning the data.

**Exploratory Findings and Results**

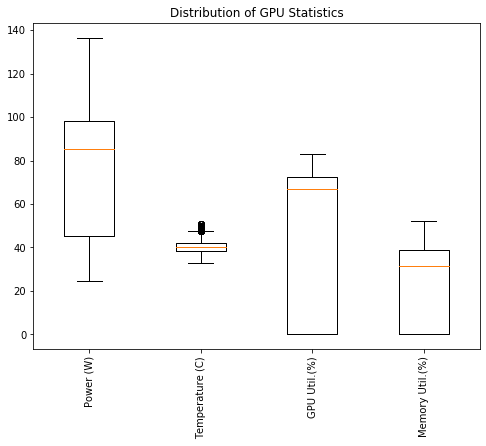
The exploratory findings have shown that the rendering task is one of the most important and time-consuming tasks in the terapixel rendering process, as it topped the execution times and used the GPUs the most.



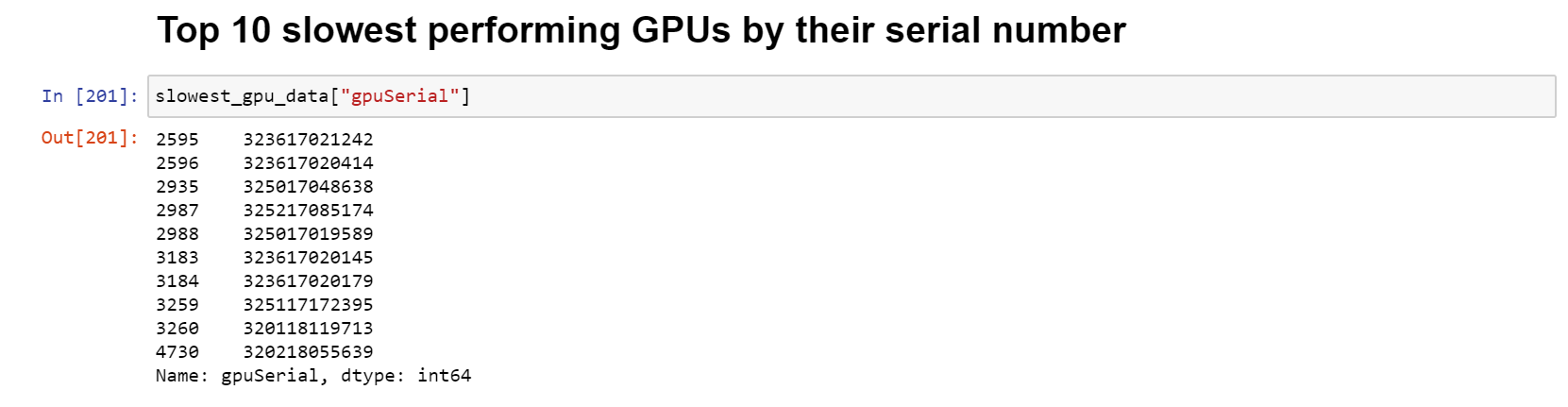
On an average around 8-12 task were assigned by every GPU.



The power consumption by the GPUs had an average of 80%, while the memory consumption and temperature were the least as in around 40C. The GPU utilisation was around 70% which seems to be decent enough



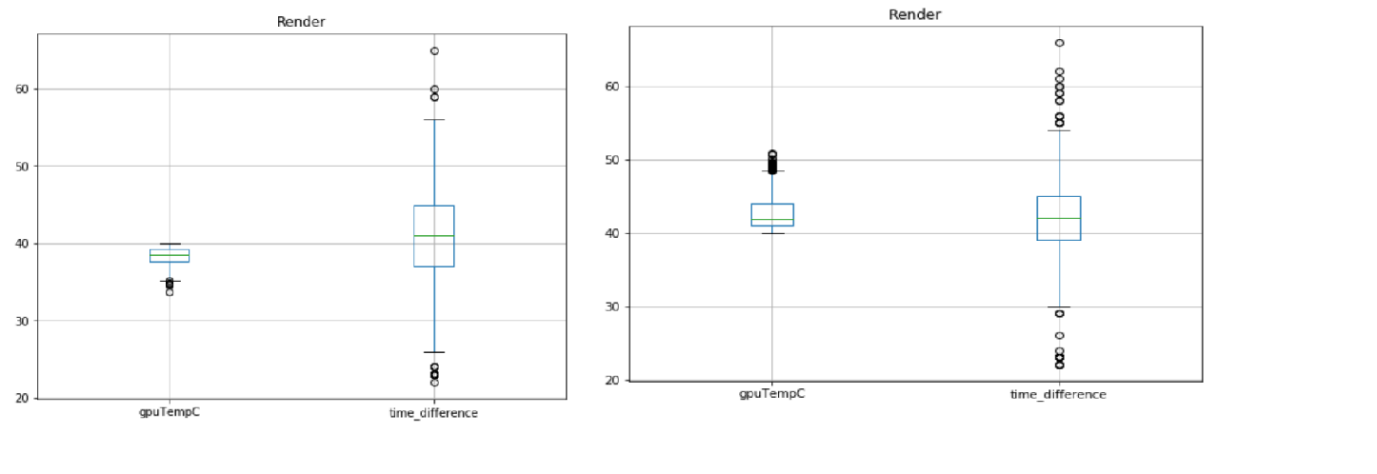
Around 10 GPUs had by far the worst performance.



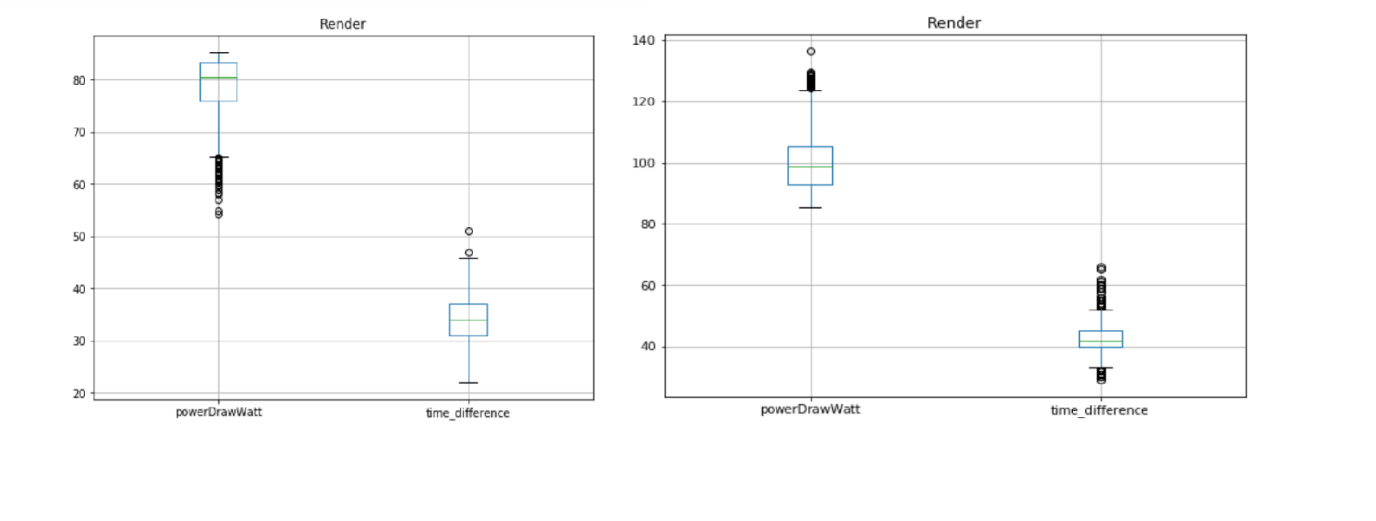
The project has found that while the GPU itself was well utilised, the memory of the GPU of itself might have not been as well utilised showing lower usages even during rendering.

It can be seen below that with increase in temperatures and power consumption the task time taken by the GPUs were also higher. This is not good for optimized results. Ideally, we would want the GPUs to consume less power and work more efficiently with lesser temperatures. The average temperature of the GPU was 40c and the average power consumed was 85W. Below are the results of the GPUs render time and how they performed below and above their respective median power drawn and temperatures.

Interplay between temperature and time difference(difference between stop and start time of a task)

****

Interplay between power consumed and time difference.



**Evaluation and Future work**

The project seems to have covered a lot on the performance of the GPUs, but they have not mentioned more details regarding memory usage, task analysis. These data if made available could have helped in more in depth analysis of the performance of the terapixel rendering with respect to memory statistics, scaling etc. This can help in variety of analysis and improve the performance based on these analyses which can be considered in future projects.

It can be understood that latest tools like sphinx document generator is very useful. Using virtual environment to build a project is much useful.