

Steps for the data Analysis

• Data Inspection

- At first, I tried to understand the data. The data Consists of around 15 columns. From these columns I found “Most recent unassign tasks” columns to be unmeaningful and I removed that column.
- Then I tried to understand the columns essential for measuring performance of the researches according to benchmarks points and the time taken to complete.

• Data Cleaning

The Data cleaning is one of the most important parts in any Data Analysis. I have done several data cleaning steps -:

- There were lot of empty values in the closed data and Workflow Priority columns so in that I have replaced the empty values with the "uncomplete" and "Null" Respectively.
- The Data format was uneven and so I have corrected the date format for all the "Date" data type Column.
- Then I splitted the time from the date columns and removed it as it only adds to the complexity of the dataset.

• Data Analysis to find the Researchers Performance Report

- I have created the detailed dashboard which provides the idea about the top 10 employees with the highest benchmarks points, Top 10 research groups with the highest benchmarks points while comparing it with the total secondary researchers in equivalent research groups and the total time taken by each research group which implies it productivity.
- I have also tried to figure out about the top 10 research group with the highest workflow and their respective workflow Priorities.
- I have also tried to categorise the status of workflow according to different region.

• Finding from the report

- My analysis reports that Researcher 56 has the highest benchmark point of 19897. The analysis suggests that among the top 10 research group according to the

benchmark points, some research group like " MUM - PC - ALL - Swing " has good benchmarks points but this is due to the high Number of Secondary Researches. Thus, the overall productivity is lower.

- Some Research groups like "MUM - PE - CL " has higher benchmarks points despite the lower number of Secondary researchers which shows it has a higher productivity level.
- Comparing according to the region as a whole it shows that North America (1st) and Europe (2nd) has the highest number of "Closed" workflow status. While the analysis suggest that India has higher number of “Pend Survey” then “closed”.
- On analysing the project requirements, we found that " MUM - F - SEC " have the highest count of workflow priority from which project categorised as workflow priority as "High" is the highest.
- For full Data Analysis Report please click on the Link for the [Google Sheets](#) or the Excel file (Detailed) attached in the mail.