To complete the work for the course, I first of all cleaned up both of the data tables where the results were noted, in order to attain better clarity and comparison potential between the different tasks, but also the different models, as I added the results of two new models.

I also rechecked the accuracy first assigned, to make sure it made as much sense as possible, as well as writing the answer I thought would've been right. This then made it easier to rate the accuracy of the other models as I imputed the results for each task.

The next step was to test the Claude model, with the same prompt and tasks as I had previously done using ChatGPT. The work was a bit more complicated as Claude, in its free version, is limited to a certain number of messages per few hours, meaning I would often have to stop and come back to complete the tasks I wanted to do.

I then retried all of the tasks and prompts using the Mistral's language model, where I discovered that the model itself offered three different levels of work. I then decided to also compare these three different levels on each task and see how they each would compare to each other, but also compared to Claude and ChatGPT.

After that, I asked the three models to categorize all of the text snippets according to their own category, to see how different the results would end up being.

To finish, I asked the three models, using the text snippets I sent, to resume the general sentiment of the player base towards the game. They all gave similar answers to this part.

To complete-, I wrote a small comparison, or analysis, of the results I opened from the three models, but also from the three different "levels" available from mistral, to resume the results that it is possible to see in the tables.

I also made a small poster to present the project at the DH Con happening in Mai.