Data Processing Problem

Think of this challenge as an opportunity to show us what "good" looks like to you. We'd like to see how you approach working with data from different sources and how you go about extracting insights from it. To make it a bit of fun, we chose two Star Wars-related data sources: the **Star Wars API (SWAPI)** and a **movie scripts dataset** from Kaggle.

- Star Wars API: https://swapi.dev/
- Scripts data set in attached in the zip archive

You don't need to spend more than a couple of hours on this. If you don't get through everything, that's totally fine—unfinished submissions are welcome. We're more interested in seeing your thought process, how you approach the problem, and what you prioritise.

The Tasks

There are three tasks for you to tackle:

1. **Proportion of Human vs Non-Human Characters** – Use the provided sources to calculate the overall proportion of human and non-human characters in the Star Wars universe, and what the proportions are in each movie.
2. **Planets and Screen Time** – Counting lines of dialogue or words spoken — whichever works best for you – rank the home planets of characters according to the characters' total time on screen, based on the film scripts. The top 5 planets are enough.
3. **Visualizing Screen Time** – Finally, we'd love to see how the "planet of origin screen time" changes over the course of the movies. Create an appropriate visualisation that shows the distribution of screen time for characters' planets of origin during the films. Feel free to use any type of chart and preprocessing that you think best shows this evolution with reasonable clarity.

What to Submit

We're interested in your results, but just as important is how you got there. Share your work in a GitHub repository, including your code (Scripts, SQL, notebooks, etc.), any data or outputs you generated, any validation of the code, etc. Once again, think of it as an opportunity to show us how you'd submit real work.

That's it! If you have any questions along the way, feel free to reach out.