

## Assignment – Data Science

Welcome to the Audible Data Science technical challenge! This challenge is designed to give you a feel of what working on a sample project might look like at Audible.

Please complete this assignment within 3 calendar days of receiving this. Email your Jupyter notebook to Alex Reynolds ([jaredre@audible.com](mailto:jaredre@audible.com)) and Jeremy Bradshaw ([jbradsh@audible.com](mailto:jbradsh@audible.com)).

### Movie Rating Prediction

Use the publicly available [IMDB Datasets](#) to build a model that predicts a movie's average rating.

Please document your approach and present your results in the Jupyter notebook. Make sure your code is well-organized so that we can follow your modeling process (e.g., from data preprocessing and feature selection to model building and evaluation). Your final model does not need to be overly complex, but you should be prepared to explain why you chose that model, and what you would experiment with if you had additional time and resources (and why).

Do not submit the data itself, but please follow the folder structure below. Make sure that we can run your notebook start to finish.

```
.
├── data
│   ├── title.ratings.tsv.gz
│   └── ...
└── <your_notebook>.ipynb
```

Once you complete the assignment, colleagues from our Data Science teams will review your work and we will follow up with next steps.

### Hints!

- We are most interested in getting a closer look at your thinking process and how you would apply your knowledge and experience to real-world use cases.
- Make as many assumptions as necessary as long as you document them. We'd prefer to have you elaborate based on false premises, rather than not being able to dive deep on your thinking process at all.
- Some example questions you may want to consider: Who would use your model? What might they find surprising? How confident would you be in your model predictions? Etc.

**Many thanks!**

Your Audible Recruiting Team