



## DATA SCIENTIST CHALLENGE

---

Welcome to the RappiPay challenge for **Data Scientists**. From the interview we had, we can already tell that you're a very knowledgeable professional that we would love to call teammate. In this challenge, we are looking for ownership: ownership of the decisions, ownership of the data, and ownership of the business, hence imagine you're already a crucial part of the team and that many business decisions rely on the input you provide. Let's get started...

**[1]** Clone the following repo:

<https://bitbucket.org/carlwhandlin/rpp-ds-challenge/src/master/>

### YOUR TASK IS TO:

- Explore the data and present a thorough analysis.
- Present a categorization of the different customers as you find fit.
- Create a model that is capable of detecting fraudsters. Put this model into production (could be locally or preferably in the cloud).
- Explain how good your model is, which are the most important features, and how the values of the variables influence in the model.
- Explain the trade-off of using your model to detect fraudsters.
- Push the project to a repo and share this with us. Don't forget to include instructions and the packages to make the system replicable.
- Share your most valuable insights of the data and the evidence of the model in production.

You won't be evaluated by how good the model performs. Don't waste your time there.

The Rappi team and I wish you the best of lucks.

---

Carl W. Handlin

**Head of Data Science, RappiPay.**

carl.wallace@rappi.com