

Assignment for Data Science Summer Intern

We are delighted to continue the recruitment process with you. Now it is your time to shine. We've prepared a task for you so that you can prove your skills and to let us understand how you are approaching problems in general. Good luck.

Exercise

In this exercise you get to look at a piece of sales data and train a small predictive model on it. The idea is not to use a lot of time on this, or to polish every corner to the max, but to demonstrate that you understand the basic concepts and can argue about your decisions.

Load the data

The orders are stored in a JSON list, one object per order. All of the orders are in the same city.

Fields

- **distance:** the distance from the restaurant to the customer
- **preorder:** whether the order has been made in advance or for immediate delivery
- **timestamp:** UNIX timestamp of the order (millisecond resolution)
- **tz:** timezone of the timestamp
- **venue_tags:** list of strings that describe the food types that the venue sells
- **weather:** string denoting the general weather conditions that day



Take a look

Visualise the data, give some descriptive statistics. Would it be useful to transform some of the variables, create new ones or drop some? Do you see any trends or patterns in food ordering habits regarding the kinds of food people order?

Train

Train a model to predict how many orders we will get. Can we predict it for the next hour? The next day? The next week? What kind of restrictions or assumptions are there? You can choose the approach.

Discuss

How do you measure the trained model? In case you tried different ways to model the problem, how can you compare the performance?

How would you improve the model or data processing? Would there be some extra information you would like to have for the task?

If you gained some interesting insight on the side, feel free to expand on it.

Return your finished assignment to laura@wolt.com by 25.2.2019.

