# Predicting Delivery Time for an On-Demand Shopping and Delivery Service

The goal of this work was to predict the total time of a delivery process for an on-demand shopping and delivery service. Four tables of data were provided, one describing the products of an order, one describing the orders, one describing the shoppers and one describing the stores.

The main tools used were Python 3.5 and Jupyter Notebooks. The main libraries used were:

- Pandas
- Matplotlib
- Seaborn
- Scikit Learn
- TensorFlow
- XGBoost

The work comprises three parts, each one of them developed in separated Jupyter Notebooks.

In part I, Data Exploration and Visualization, the data provided was deeply studied using descriptive statistics and visualization tools. The objective was to understand the data and to get insights on how variables relates and behave.

In part II, Data Processing and Transformation, the dataset for training and testing was generated. In this part, new variables were created, transformations to the data were applied, NA values were processed and joins and merges of data-tables were done.

Finally, in part III, Model Training, Selection and Predictions, three machine learning models were trained and tested, a Linear Regression model, a Gradient Boosting model and a Neural Network of three hiden layers with twenty nodes each. Then, the one with the best performance was used to make predictions on the orders required.