## Ctrl + shift + p vscode

## Project 1: Predict total sales for every product and store in the next month.

In this project you will work with a dataset of a store chain in Russia consisting of daily sales data.

For successful completion, you are required to do the following tasks:

* Data Preprocessing
* Data Exploration
* Feature Selection and Correlation
* Splitting Testing and Training Data
* Modeling
* Training and Evaluation
* Model Evaluation

Dataset format: CSV Files

Complete Dataset description will be provided.

ML Models: Open choice to choose any ML model.

## File descriptions

* sales\_train.csv - the training set. Daily historical data from January 2013 to October 2015.
* test.csv - the test set. **You need to forecast the sales for these shops and products for November 2015.**
* sample\_submission.csv - a sample submission file in the correct format.
* items.csv - supplemental information about the items/products.
* item\_categories.csv - supplemental information about the items categories.
* shops.csv- supplemental information about the shops.

## Data fields

* ID - an Id that represents a (Shop, Item) tuple within the test set
* shop\_id - unique identifier of a shop
* item\_id - unique identifier of a product
* item\_category\_id - unique identifier of item category
* item\_cnt\_day - number of products sold. You are predicting a monthly amount of this measure
* item\_price - current price of an item
* date - date in format dd/mm/yyyy
* date\_block\_num - a consecutive month number, used for convenience. January 2013 is 0, February 2013 is 1,..., October 2015 is 33
* item\_name - name of item
* shop\_name - name of shop
* item\_category\_name - name of item category

**Grading criteria**

We follow the guideline for bachelor theses in Engineering (<https://www.uhr.no/_f/p1/i2803bcbe-1f46-4822-8a70-f0e52f5a6a8f/eng_grade_descriptions_and_assessor_information_bachelorthesis_uhr_logo.pdf>)

**Report template**

General sections from a typical technical document are OK, a good example:

1. Introduction
   1. Project motivation
   2. Project scope
   3. What you have achieved
2. Background

Background information needed to understand the technical details in the rest of the report

1. (Your implementation, e.g., Object Dectection for Security Scanners)
2. Experiments
   1. Datasets
   2. Results
   3. Analysis, comparisons, and discussions
3. Conclusion

**Oral session (after the delivery of project reports)**

We will have a 5-minute oral Q&A session with each project group, no presentation needed. Mainly to have discussions on some technical and implementation details.