HSE, Data Science Case Study , Fall 2025

HW #3

Analytics in Retail

alexromsput@gmail.com

Organization Info

Submitting results:

- Submit homework via the Yandex form as a link to your GitHub, where all files and code (Python) will be.
- GitHub must be open and the code must be working, without errors.
- Name the repository using the template (HW_3_2025-_).
- Link to the Yandex form: https://forms.yandex.ru/u/68eece24505690c23425594c

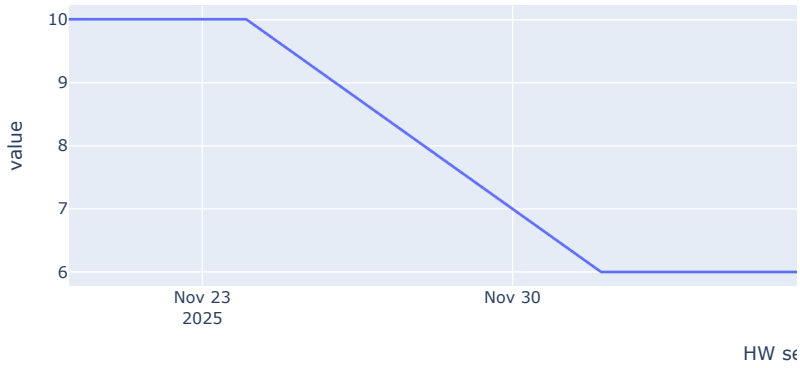
⚡ Date of issue: October 26, 2025

- ⏰ Soft deadline - 11:59 PM MSK on November 23, 2025
- ⏰ Hard deadline - 11:59 PM MSK on December 30, 2025

After the soft deadline, the grade will be deducted at a rate of 0.5 points per day.

```
import pandas as pd
import numpy as np
import datetime
pd.options.plotting.backend = "plotly"
deadline_dttm = pd.to_datetime('2025-11-24 00:00:00')

pd.DataFrame(data = [[deadline_dttm+datetime.timedelta(days=x)
                        , min(10, max(6, (10-x/2)))] for x in np.linspace(-4,30,35000)],
                        columns = ['HW sending date', 'Max mark']).set_index('HW sending date').\
  plot().update_layout(height=350, width=1350).show()
```



Kaggle contest (100%)

- Participate in kaggle contest https://www.kaggle.com/t/71cb778738a14d7fb977353bb9eee132
- it is allowed to use any python or R packages
- use starting code from here
- 2 benchmarks:
  - you'll get $mark \in (6, 8]$ of 10 if your approach outperforms benchmark2
  - you'll get $mark \in (8, 10]$ of 10 if your approach outperforms benchmark1
  - final $mark$ will be derived based on leaderboard
  - you have to send code with the final solution to Yandex form https://forms.yandex.ru/u/68eece24505690c23425594c (otherwise final mark will be 0)

Additional material for homework:

- Example of data analysis with segmentation and profiling discussed in the seminar: https://colab.research.google.com/drive/15nzMVBa5twjAkAVCf854wkSA9WhanaW5?usp=sharing;
- An example of a research paper description in presentation format - a cheat sheet for homework is provided in the folder with options https://disk.yandex.ru/d/i0z880-cKcTSzQ;
- Link to a video description of the solution to homework 1 (Added on the evening of October 16, 2025);

GUI:

- https://desktop.github.com/
- https://git-scm.com/
- https://tortoisegit.org/

Useful links on each topic:

- Using ML for Demand Forecasting https://github.com/aromanenko/DSCS/blob/main/DSCS_4_Seminar_Analtics_in_SC_v1.ipynb
- Dl for Demand Dorecating https://github.com/aromanenko/DLDF/tree/master
- Demand Forecastin in Retail: https://github.com/aromanenko/DSCS/blob/main/DSCS_5_Lecture_Analtics%20in%20SC_v1.pdf