

DAT 562 TEXT MINING

PROF. YULIA NEVSKAYA

- Lab 5: Sentiment Analysis
- Office Hours (Zoom Link is on Canvas):
 - Wednesdays 9:30-10:30am CT
 - Fridays 12:30-1:30pm CT



Reminder: Project Pitch

- Project Scope is **due December 2, 11:59pm (St. Louis time)**
 - 2-3 "McKinsey style" slides + title slide
 - Include photos of teammates (+ indicate a presenter if attending remotely)
 - See Project Guidelines for more details (on Canvas)
 - Check out a guide to the "McKinsey style" on Canvas
 - No late submissions
- Next lab:
 - Use your Project Scope slides to pitch your project
 - No updates to the slides after submission on Canvas

Next Lab: Rules for the Pitch

- Timing for Presentation is strict: **Max is 7 mins, Min is 6 mins**
 - Make your pitch perfect: Rehearse, rehearse, rehearse!
- Q&A after each presentation
 - **Everyone can** ask questions
 - **Reviewers must** ask questions and get priority
- Nominate **1 team member** to present
- I'll announce the order for teams' pitches on Canvas the day before
- Questions?

Peer Pitch Review

- Each student **reviews 1 pitch** delivered by another team:
 - Your review will be shared with the reviewed team
 - You will see which team you will be reviewing on Canvas, on Saturday
 - Look at the slides of the team you review before next lab
- Submit a 1-page review by **December 9, 11:59pm, CT:**
 - *What do you like about the project?*
 - *Possible issues: how realistic are the project goals? Are the goals managerially relevant? Are methods well aligned with the goals? Any selection issues with the data?*
 - *How can the pitched project be improved?*
- Use your opportunity to **ask questions during Q&A**
- See Project Guidelines for more details (on Canvas)

The Plan for Today

- We'll be doing **sentiment analysis** in Python
- In particular, you will:
 - Perform VADER lexicon-based sentiment analysis (unsupervised machine learning)
 - Fine-tune the lexicon-based sentiment analyzer
 - Train your own sentiment analyzer (supervised machine learning)

What You Need to Do Today

- Go to Canvas / Module 5 / During Lab to get the Assignment
 - Download 2 Jupyter notebooks + **2 data files** and put them into the same directory(!) on your machine
- **YOUR TASKS FOR TODAY:**
 - 1) Do all the **EXERCISES** in the script
 - 2) Submit individually the Jupyter notebook with your results and answers on Canvas by the end of the session

Reminders

- The Project Scope is due December 2, 11:59pm CT
- Office hours in person or on Zoom (link in the Syllabus):
 - Wednesdays 9:30-10:30am CT
 - Fridays 12:30-1:30pm CT

Jupyter.org/try (Try JupyterLab)

[Install](#) [About Us](#) [Community](#) [Documentation](#) [NBViewer](#)

Try Classic Notebook



A tutorial introducing basic features of Jupyter notebooks and the IPython kernel using the classic Jupyter Notebook interface.

Try JupyterLab



JupyterLab is the new interface for Jupyter notebooks and is ready for general use. Give it a try!

Try Jupyter with Julia



A basic example of using Jupyter with Julia.

Try Jupyter with R



A basic example of using Jupyter with R.

Try Jupyter with C++



A basic example of using Jupyter with C++

Try Jupyter with Scheme



Explore the Calysto Scheme programming language, featuring integration with Python