

# 189G Week 1 Discussion

# Yuyi Li

- Email: [nylli@ucdavis.edu](mailto:nylli@ucdavis.edu)
- Office hour: M 5-6 pm, W 5-6 pm, Kemper 55
- Interactive Grading time: W 2-3 pm, Kemper 3106
- Group forming due this week!
- Email me if you want to group with/without someone specifically!

# Overview

- OMSI system testing
- New feature introduction: chat
- R for basic linear algebra
- 'Compas' dataset
- 'German credit' dataset

# OMSI(for HW & Quiz)

TL;DR:

Step 1: Download OMSI from <https://github.com/matloff/omsi>

Step 2: Connect to Library VPN, because we are

Step 3: Go to the OMSI directory, find 'OmsiGui.py'

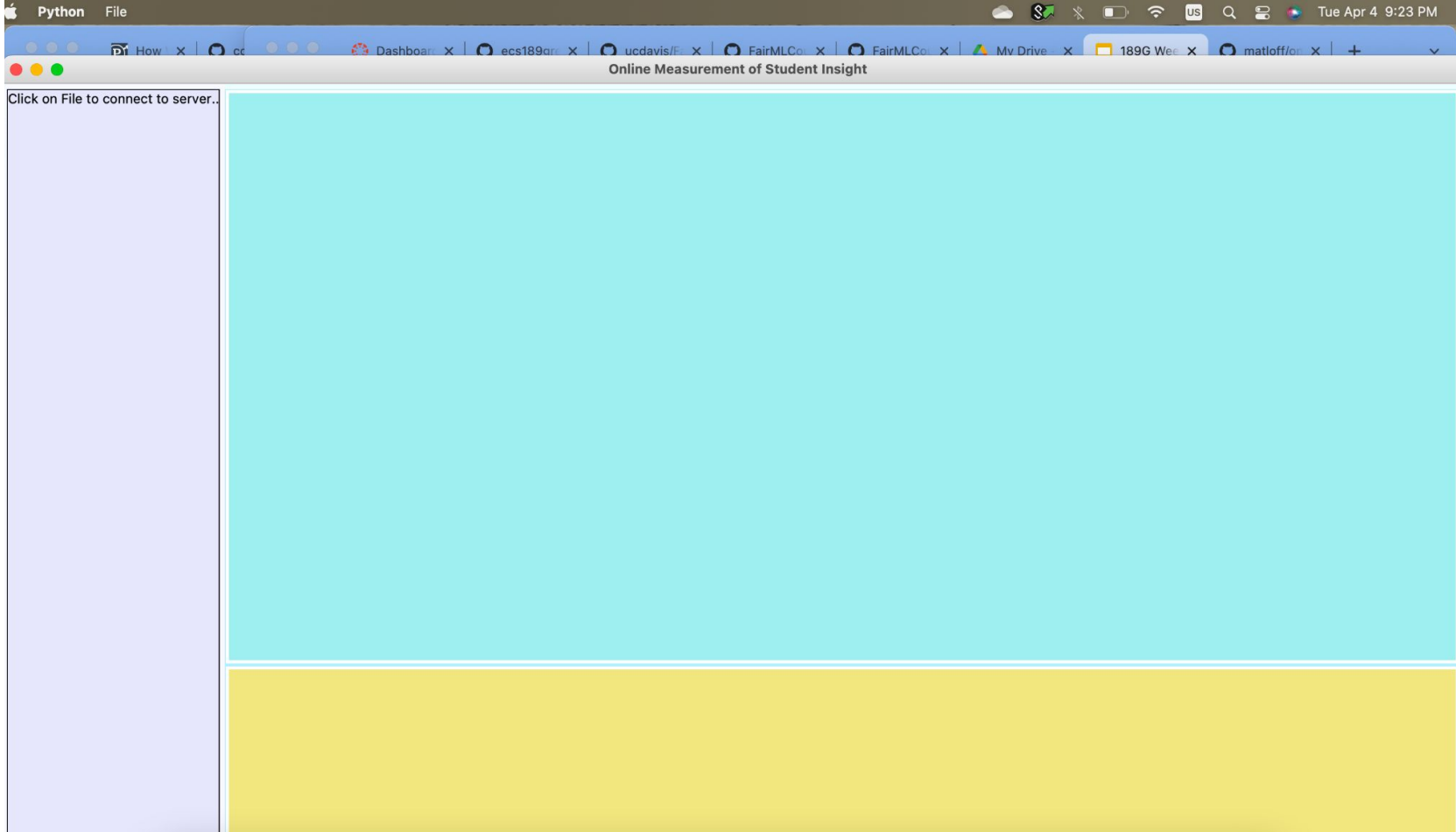
Step 4: Run 'OmsiGui.py' on your terminal with the following command line script:

```
python3 OmsiGui.py (localhost) (port #) (your UCD email)  
(quiz name)
```

**E.g.:**

```
python3 OmsiGui.py pc11.cs.ucdavis.edu 4000 nylli@ucdavis.edu  
quiz0_sq23
```

By now, you should have seen something like this:



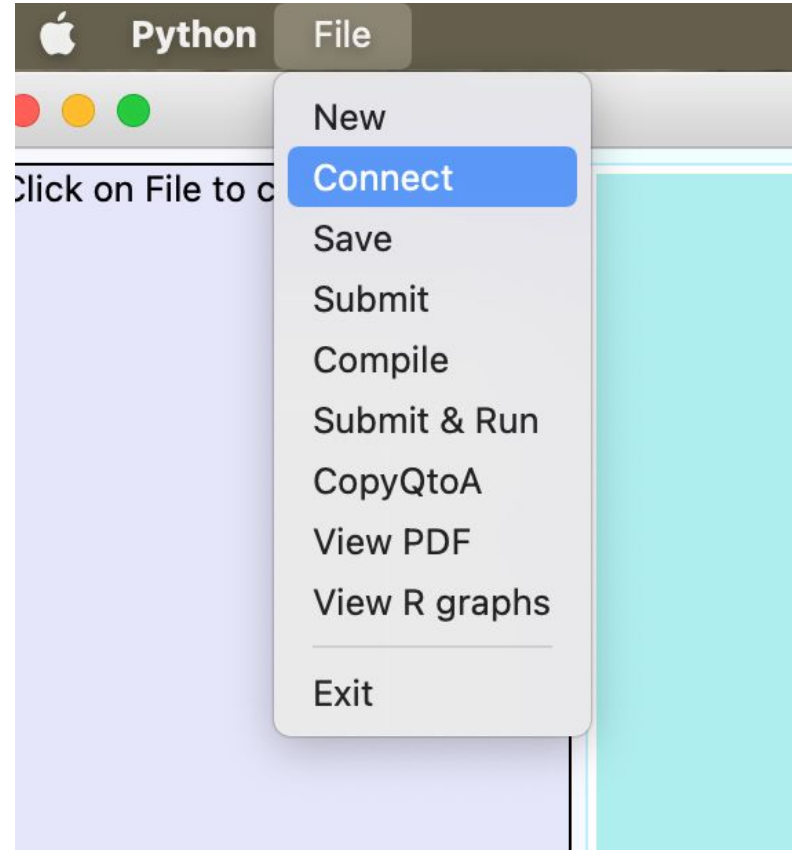
# OMSI(Cont'I)

Step 5: on the Top right corner, hit 'File' and select 'Connect' in the Drop Down menu bar

Step 6: Now you should see the problems on the left hand side. Select whichever problem you want to answer first

Step 7: Every time you finish answering a problem, select 'Save' in the Drop Down menu bar

Step 8: After finishing answering all questions, select 'Submit&Run' and close the window.





**nylli@ucdavis.edu successfully  
submitted omsi\_answer1.R**

**OK**

# R for basic linear algebra

- Matrix Multiplication
- Find the determinant of a matrix
- Inverse of Matrix
- Find the eigenvalue and eigenvectors of a matrix
- Etc
- Full link: <https://github.com/matloff/fasteR>



# Background of 'compas' dataset

- Correctional Offender Management Profiling for Alternative Sanctions
- From Broward County, Florida
- “...using algorithms to assess a criminal defendant’s likelihood of becoming a recidivist – a term used to describe criminals who reoffend. ”
- <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>

# Background of 'German credit'

- The dataset includes 1,000 examples and 20 input variables, 7 of which are numerical (integer) and 13 are categorical.
- Classifying bank customers as to whether they should receive a loan or not. Giving a loan to a bad customer marked as a good customer results in a greater cost to the bank than denying a loan to a good customer marked as a bad customer.
- This requires careful selection of a performance metric that both promotes minimizing misclassification errors in general, and favors minimizing one type of misclassification error over another.

## Coming up next...

- No quiz next week
- Some R tricks for HW

qeML::checkPkgLoaded

Link below:

<https://github.com/matloff/qeML>