

PSTAT 131/231: Introduction to Statistical Machine Learning

Guo Yu

Department of Statistics and Applied Probability
UCSB

Course Pages

Gauchospace

Course hub page: lecture slides, quizzes, homework assignments ...

Piazza for all course related Q&A

<https://piazza.com/ucsb/fall2021/pstat131231>

Office Hours

Guo Yu (Instructor)

Tu/Th 5:15 - 6:15 PM

Hanmo Li (TA)

M 5:00 - 7:00 PM

Chau Tran (TA)

F 3:30 - 5:30 PM

Eric Bletcher (ULA)

M 1:00 - 2:00 PM

Michael La (ULA)

TBD

**All office hours will be held synchronously via Zoom meeting
(links available on Gauchospace)**

Guo's OH starts today!

References

Introduction to Statistical Learning, with Applications in R, James et al

aka **ISL**, for all students

Element of Statistical Learning, Hastie et al

aka **ESL**, only for 231 student

R for Data Science, Grolemund and Wickham

aka **R4DS**, will not be covered in class, only for self learning

All freely available online!

Grading

40% — Homework (4 hw assignments, due ~ every 2 weeks)

25% — Midterm Exam (Nov 4, 3:30 - 4:45 PM during class)

10% — Quizzes (on Gauchospace)

25% — Final project (will be out mid Nov, due 7 PM, Dec 9th)

Homework

R coding + math derivation/proof

Typically you will have 2 weeks to complete the homework

You are encouraged to work with a partner

Late homework will receive 10% point deduction

Homework will not be accepted more than 24 hrs late

Lectures vs Lab Sessions

Lectures focus on concepts/methods

Lab sessions focus on data analysis / method implementation

131 vs 231

All lecture materials will be accessible to all students

231 students have extra homework, midterm, project questions

131 and 231 students are graded on different scales

R, Rstudio, and Rmarkdown

R is a user-friendly programming language for statistical analysis

RStudio is an integrated development environment (IDE) for **R**

R markdown (Rmd) generates reproducible document with **R**

- Syntax is very simple
- Integrated with **LaTeX**, easy to write math formulas
- Example: **homework-template.Rmd** (on GauchoSpace)

R, Rstudio, and Rmarkdown

R is a user-friendly programming language for statistical analysis

RStudio is an integrated development environment (IDE) for **R**

R markdown (Rmd) generates reproducible document with **R**

- Syntax is very simple
- Integrated with **LaTeX**, easy to write math formulas
- Example: **homework-template.Rmd** (on GauchoSpace)

Start setting up the R working environment!

Deliverables

Homework and final project submission format:

- **Electric submission via GauchoSpace**
- **A zip file containing:**
 - **R markdown code**
 - **Generated pdf file**
 - **Any supplementary file**

Deliverables

Homework and final project submission format:

- **Electric submission via GauchoSpace**
- **A zip file containing:**
 - **R markdown code**
 - **Generated pdf file**
 - **Any supplementary file**

We should be able to run your RMD code to obtain the same pdf file as your submission

Next...

Join Piazza!

Make sure you can knit **homework-template.Rmd**

Lab session starts next week!

Read R4DS Chapter 1&2 (very short), and Chapter 27 (on Rmd)