In-Class Announcements
(Slides also available on course homepage)

Duen Horng (Polo) Chau
Assistant Professor
Associate Director, MS Analytics
Georgia Tech

Partly based on materials by
Professors Guy Lebanon, Jeffrey Heer, John Stasko, Christos Faloutsos
Thu, 1/12
How to address Polo?

**Grammatically correct**

Prof. Chau

Dr. Chau

**Grammatically incorrect, but popular**

Prof. Polo

Dr. Polo
Course Homepage
For syllabus, HWs, projects, datasets, etc.

Google “cse6242”
polo.club.gatech.edu/cse6242/2017spring

All students must first review prerequisites & course expectation.

CSE6242 / CX4242, Spring 2017
Data and Visual Analytics
Georga Tech, College of Computing
Join Piazza ASAP

goo.gl/b0TkEP

Announcements and Discussion

We use Piazza for announcements and discussion.

Everyone must join this class’s Piazza, at https://piazza.com/class/ixpgu1xccuo47d.

Double check that you are joining the right Piazza!

When you have questions about class, homework, project, etc., post your questions there. Our teaching staff and your fellow classmates will help answer them quickly. You can also use Pizza to find project teammates.

T-square will only be used for submission of assignments and projects.

While we welcome everyone to share their experiences in tackling issues and helping each other out, but please do not post your answers, as that may affect the learning experience.
Important to join Piazza because...

The fastest way to get help with homework assignments is to post your questions on Piazza. If you prefer that your question addresses to only our TAs and the instructor, you can use the private post feature (i.e., check the "Individual Students(s) / Instructors(s)" radio box).
Important to join Piazza because...

- Polo will announces events related to this class and data science in general
- Distinguished lectures
- Seminars
- Hackathons (free food, prizes)
- Company recruitment events (free food, swag)
Policies

Collaborating on homework
Late submission policy
While collaboration is allowed for homework assignments, each student **must** write up their own answers. All GT students must observe the honor code. Any suspected plagiarism and academic misconduct will be reported and directly handled by the Office of Student Integrity (OSI).
WARNING
You’ll be writing a lot of code

Q: Is it OK to copy and use code found on the web?
A: No

Q: Why?
A: Here’s why…
Do not plagiarize!

• Using code as reference does not mean copying and pasting that code. Nor does that mean copying in a block of code and then modifying parts of it.

• If you want to use some code for reference, you should go over it, understand what it is doing, and then try to accomplish what it is trying to do using your own code. And it’s a good practice to cite the sources (e.g., as part of your code comments).

• The analogy is like how you would write an essay or a speech. You can get inspirations from others, but you should use your own words, otherwise it will be considered plagiarism. As I mentioned in class, and in the beginning of every homework, plagiarism can lead to heavy consequences.

Late Submissions Policy

- **Homework:** each student has **4 slip days** total. No questions asked.
- **Project:** each team has **3 slip days** total. No questions asked. Slip days may not be used on in-class activities (e.g., proposal presentation, poster presentation, etc.).
- To use slip days, **specify the number of days you have used in the textbox on T-Square** (when you submit your work).
- **Each slip day equals 24 hours.** E.g., if a submission is late for 30 hours, that counts as 2 slip days.
- After all slip days are used up, **5% deduction for every 24 hours of delay.** (e.g., 5 points for a 100-point homework)
- **We will not consider late submission of any missing parts** of an homework assignment or project deliverable. To make sure you have submitted everything, download your submitted files to double check.
- No penalties for medical reasons or emergencies. You **must** submit a doctor's note or an official letter explaining the emergency.
Distance Learning Sections (Q & Q3)

A standard 3-day lag applies to all homework and project deliverables. For project presentation, a group that has DL student member can choose to:

1. Present in class without 3-day lag; or
2. Submit a video presentation with 3-day lag (e.g., screen capture)
Prerequisites & Expectation
For both CSE 6242 (grad) and CX 4242 (undergrad)

Students are expected to complete significant programming assignments (homework, project) that may involve higher-level languages or scripting (e.g., Java, R, Matlab, Python, C++, etc.). Some assignments may involve web programming and D3 (e.g., Javascript, CSS).

You are expected to quickly learn many new things. For example, an assignment on Hadoop programming may require you to learn some basic Java and Scala quickly, which should not be too challenging if you already know another high-level language like Python or C++. Please make sure you are comfortable with this.

Please take a look at the assignments (homework and project) of the previous offerings of this course, which will give you some idea about the difficulty level of the assignments.

Basic linear algebra, probability knowledge is expected.

e.g., http://poloclub.gatech.edu/cse6242/2016fall/
FREE After-class Coffee 🍩

- After each class, Polo randomly selects 5 students (+2 volunteers) for FREE after-class coffee
- Polo’s treat. You can order coffee, tea, pastries — whatever you want.
- Very casual — you can ask me ANYTHING
Tue, 1/17
Course Project

• You’re welcome to look for teammates now. Deadline: **Friday, Feb 10** (on course schedule)

• **4-6 people** in each team

• Polo suggests teaming with students in the **same section**, but OK to mix if really needed. See “Teaming” section on project description: http://poloclub.gatech.edu/cse6242/2017spring/project.html
Course Project

• Two example projects presented next week
  • Firebird: Predicting Fire Risk in Atlanta
  • PASSAGE: Safe Path Recommendation
• Q&A after presentations
Course Project

• Can we form a 7-person team?
  • I highly recommend not to. Only with my permission.
  • Larger groups are harder to run.
  • Higher coordination, motivation and intellectual costs

https://www.cmu.edu/teaching/designteach/design/instructionalstrategies/groupprojects/challenges.html
HW1

• Out this Friday

• At least 2 weeks to complete
Thu, 1/19
HW1

- Out tomorrow
- At least 2 weeks to complete
ABSTRACT

The Atlanta Fire Rescue Department (AFRD), like many municipal fire departments, actively works to reduce fire risk by inspecting commercial properties for potential hazards and fire code violations. However, AFRD’s fire inspection practices relied on tradition and intuition, with no existing data-driven process for prioritizing fire inspections or identifying new properties requiring inspection. In collaboration with AFRD, we developed the Firebird framework to help municipal fire departments identify and prioritize commercial property fire inspections, using machine learning, geocoding, and information visualization. Firebird computes fire risk scores for over 5,000 buildings in the city,
PASSAGE: A Travel Safety Assistant With Safe Path Recommendations For Pedestrians

Matthew Garvey
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
mgarvey6@gatech.edu

Meghna Natraj
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
mnatraj@gatech.edu

Nilaksh Das
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
nilakshdas@gatech.edu

Bhanu Verma
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
bhanuverma@gatech.edu

Jiaxing Su
College of Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA
jiaxingsu@gatech.edu

Abstract
Atlanta has consistently ranked as one of the most dangerous cities in America with over 2.5 million crime events recorded within the past six years. People who commute by walking are highly susceptible to crime here. To address this problem, our group has developed a mobile application, PASSAGE, that uses real-time and crime data to find "safe paths" for users. The application provides safe paths for users in Atlanta.

Authors
Matthew Garvey
Meghna Natraj
Nilaksh Das
Bhanu Verma
Jiaxing Su

ACM Reference

Figure 1: Paths recommended by PASSAGE.
Today’s random 5 students (+2 volunteers) for FREE after-class coffee ☕️
Tue, 1/24
HW1

• Out last Friday.

• Due: **Friday, Feb 3, 11:55pm eastern**

  • **START NOW! DON’T WAIT.** Come to office hours. Ask questions on Piazza.

  • Some tasks take time to run

  • (In HW3, some take a loooong time to run)
Office hours

- Announced on [Piazza](https://piazza.com/class/ixpgu1xccuo47d?cid=46)

<table>
<thead>
<tr>
<th>Day</th>
<th>Time</th>
<th>Location</th>
<th>Instructors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuesday</td>
<td>3:30-4:00 PM</td>
<td>Klaus 1324</td>
<td>Polo</td>
</tr>
<tr>
<td>Monday</td>
<td>1-2 PM</td>
<td>Klaus- Open area next to Polo's office (Klaus 1324)</td>
<td>Meghna and Fred</td>
</tr>
<tr>
<td>Wednesday</td>
<td>1-2 PM</td>
<td>CULC (3rd floor), Common area near 325</td>
<td>Bhanu and Chirag</td>
</tr>
<tr>
<td>Thursday</td>
<td>12-1 PM</td>
<td>Klaus- Open area next to Polo's office (Klaus 1324)</td>
<td>Varun and Kiran</td>
</tr>
</tbody>
</table>
Office hours ... and on **course homepage**

<table>
<thead>
<tr>
<th>Name</th>
<th>Days</th>
<th>Time</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polo Chau</td>
<td>Tue</td>
<td>3:30-4:00pm</td>
<td>Klaus 1324 + FREE after-class coffee, at Clough Starbucks</td>
</tr>
<tr>
<td>Meghna Natraj</td>
<td>Mon</td>
<td>1-2PM</td>
<td>Klaus- Open area next to Polo’s office</td>
</tr>
<tr>
<td>Head TA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fred Hohman</td>
<td>Mon</td>
<td>1-2PM</td>
<td>Klaus- Open area next to Polo’s office</td>
</tr>
<tr>
<td>Bhanu Verma</td>
<td>Wed</td>
<td>1-2PM</td>
<td>CULC (3rd floor), Common area near 325, take right from stairs, walk few steps, common area is on the left</td>
</tr>
<tr>
<td>Chirag Tailor</td>
<td>Wed</td>
<td>1-2PM</td>
<td>CULC (3rd floor), Common area near 325, take right from stairs, walk few steps, common area is on the left</td>
</tr>
<tr>
<td>Kiran Sudhir</td>
<td>Thu</td>
<td>12-1PM</td>
<td>Klaus- Open area next to Polo’s office</td>
</tr>
<tr>
<td>Varun Bezzam</td>
<td>Thu</td>
<td>12-1PM</td>
<td>Klaus- Open area next to Polo’s office</td>
</tr>
</tbody>
</table>
Class Video Recordings
Available on T-Square.
Today’s random 5 students (+2 volunteers) for FREE after-class coffee ☕
ABSTRACT

The Atlanta Fire Rescue Department (AFRD), like many municipal fire departments, actively works to reduce fire risk by inspecting commercial properties for potential hazards and fire code violations. However, AFRD’s fire inspection practices relied on tradition and intuition, with no existing data-driven process for prioritizing fire inspections or identifying new properties requiring inspection. In collaboration with AFRD, we developed the Firebird framework to help municipal fire departments identify and prioritize commercial property fire inspections, using machine learning, geocoding, and information visualization. Firebird computes fire risk scores for over 5,000 buildings in the city,
PASSAGE: A Travel Safety Assistant With Safe Path Recommendations For Pedestrians

Matthew Garvey
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
mgarvey6@gatech.edu

Meghna Natraj
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
mnatraj@gatech.edu

Nilakash Das
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
nilakashdas@gatech.edu

Bhanu Verma
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
bhanuverma@gatech.edu

Nilakshi Das
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
nilakshidas@gatech.edu

Jiaxing Su
College of Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA
jiaxingsu@gatech.edu

Abstract
Atlanta has consistently ranked as one of the most dangerous cities in America with over 2.5 million crime events recorded within the past six years. People who commute by walking are highly susceptible to crime here. To address this problem, this group has developed a mobile application, PASSAGE, that uses crime data to find "safe paths" for pedestrians in Atlanta.

Author Information
Safe Path Assistance using Pulse
ACM SIGSPATIAL 2016
H.5.2 Information Interfaces and Systems
Jiaxing Su, Nilakshi Das, Meghna Natraj, Bhanu Verma
Department of Computer Science
Georgia Institute of Technology
Atlanta, GA 30332, USA

Figure 1: Paths recommended by PASSAGE

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
Today’s random 5 students (+2 volunteers) for FREE after-class coffee ☕️