URP 6931. Introduction to Urban Analytics

# Lecture 01: Overview

Instructor: Shenhao Wang
Assistant Professor, Director of Urban AI Lab
Department of Urban and Regional Planning
University of Florida

# Overview - outline

1
Why do you need to take this class?

2
What is this class about?

3
How is this class taught?

4
Q&A

# Part 1. Why is this course?

1. Urban analytics can address challenges and leverage opportunities in cities.
2. Urban analytics formalize and generalize traditional planning approaches and your intuition.
3. Urban analytics aligns with the trends in the department, university, and society.

# Cities are facing challenges.

**Urban Mobility**

- Crowding (train, platform)
- Passengers left behind
- Over capacity operations
- Safety concerns
- System disruptions

**Other challenges**

- Climate change
- Pollution
- Economic growth
- Energy

Cities are facing opportunities

# Urban analytics provides the analytical methods to address challenges and leverage opportunities in cities

**Technology**

- Automation
- Electrification
- 5G/Connection
- Sharing Economy

**Data**

- Surveys
- Ubiquitous sensing
- Urban imagery
- Social and spatial networks

**Analytics**

1. Statistics
2. Network analysis
3. ML & DL
4. Others (causal inference, etc.)

However, traditional qualitative and quantitative planning methods also address challenges and leverage opportunities in cities,

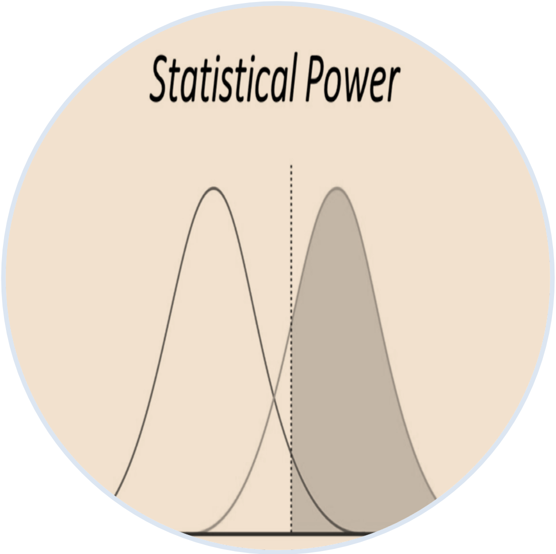
# Why urban analytics?

2. Urban analytics formalizes and generalizes traditional planning approaches (quantitative and qualitative methods) and your intuition.

# Examples

## Equity

**The racial majority groups have higher income than the minority groups.**




Formalize
- Statistics tradition - Regressions

Generalize
- Significance – Test - Confounding

## Segregation

**The high income groups are spatially segregated.**

Formalize
- Network tradition – Spatial networks

Generalize
- Nodes & edges – Spatial cluster – Community detection

## Design

**The urban designers/planners can create new landscapes to improve accessibility.**

Formalize
- Machine learning – Generative models

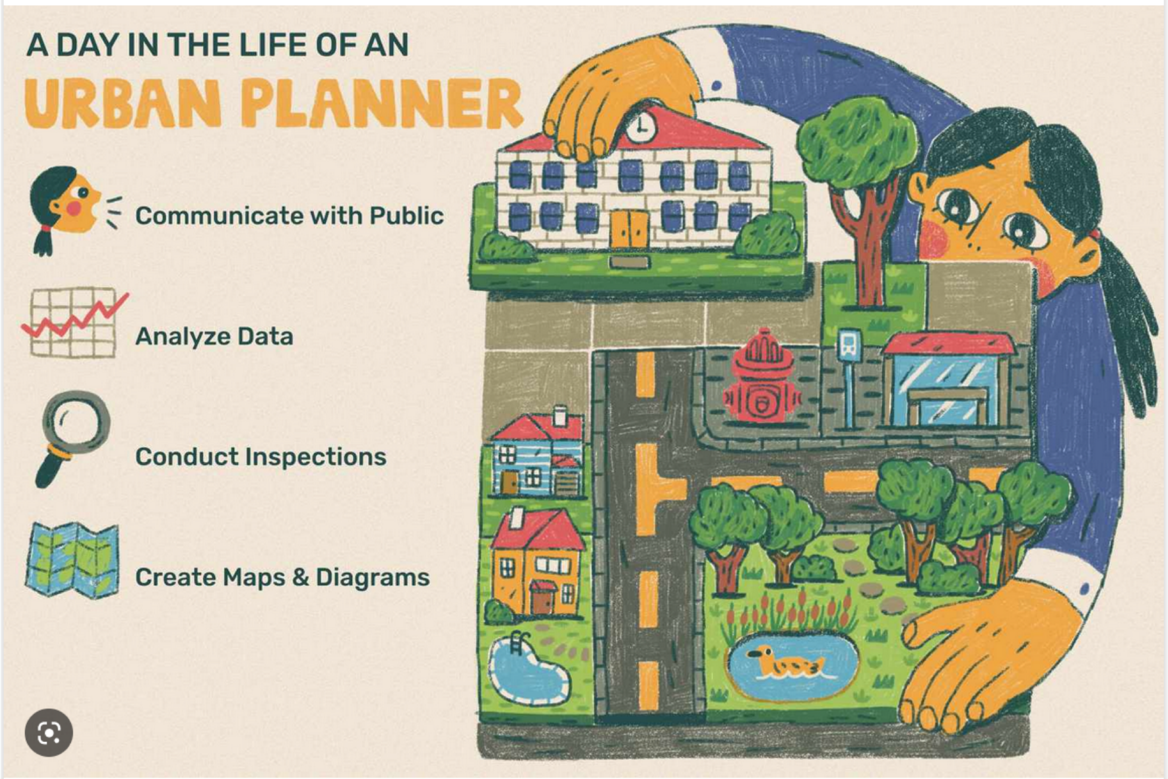
Generalize
- Latent space – Conditional GAN

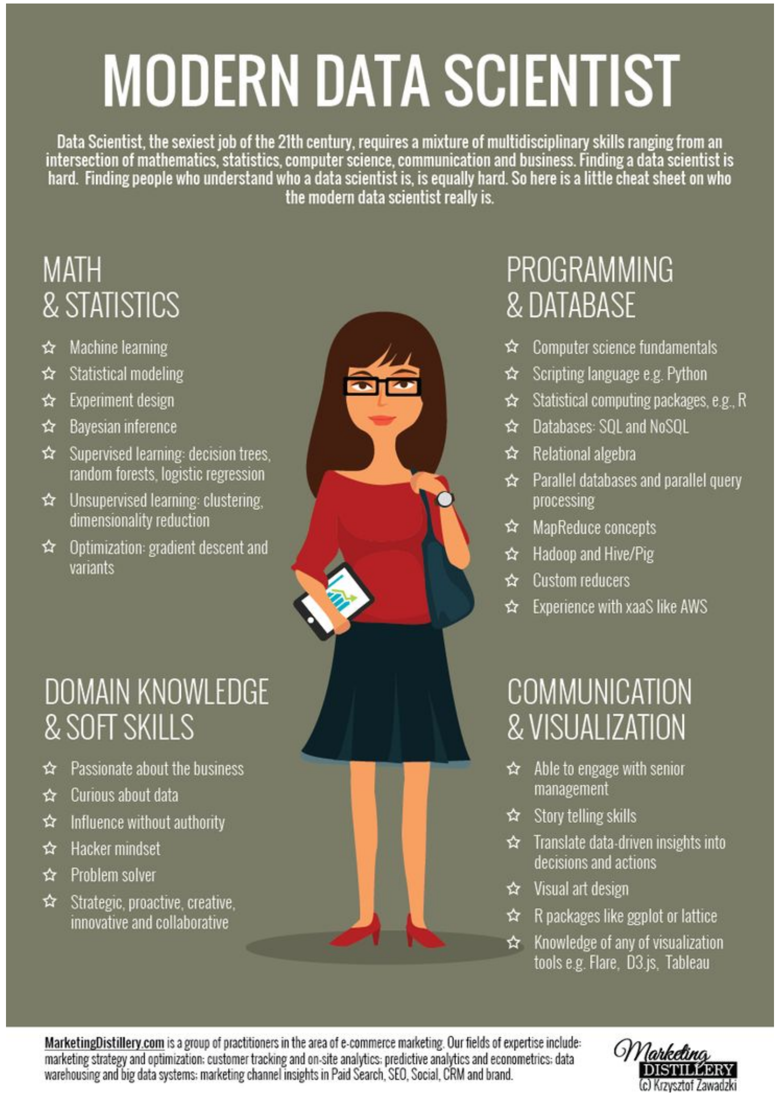
# Why urban analytics?
# Trends in department, university, and society

- **University**. UF AI initiative https://ai.ufl.edu/. "Building an AI university, transforming the future workforce"
- **Department**. Master of Urban Analytics program (Fall 2023);

| 2022Fall - 2023Spring | 2023Fall - 2024Spring | 2024Fall - 2025Spring |
| --- | --- | --- |
| 2023 Spring<br>Introduction to UA (G)<br>Shenhao Wang | 2023 Fall<br>Introduction to UA (G)<br>Yan Wang | 2024 Fall<br>Introduction to UA (G)<br>Yan Wang |
| | 2023 Fall<br>AI in Built Environment (U)<br>Shenhao Wang | 2024 Fall<br>AI in Built Environment (U)<br>Shenhao Wang |
| | 2024 Spring<br>Intermediate UA (G)<br>Shenhao Wang | 2025 Spring<br>Intermediate UA (G)<br>Shenhao Wang |
| | | 2025 Spring (potentially)<br>Advanced Urban AI (G)<br>Shenhao Wang |

# Trend in society: From traditional urban planner to urban data scientists



Individual efforts matter.

However, if you can align your individual efforts with the social trends, your achievements can be significantly amplified.

# Part 2. What is this urban analytics class about?

# Introduction to Urban Analytics

1 2

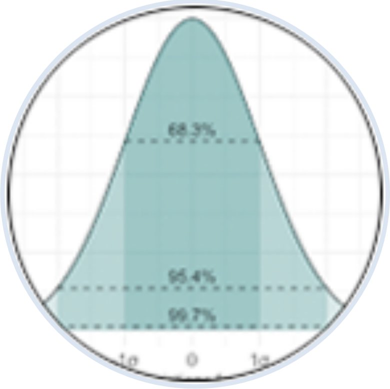
This course introduces the primary modeling paradigms to use multiple data representations to tackle the challenges in cities with an emphasis on analytical perspectives, urban applications, and computational practices .

3

# Three Primary Modeling Paradigms

## Urban Statistical Analysis

**Modeling Paradigm 1**




**IID data – Linear & logistic regressions – applications to economic and travel analysis – critiques to regressions**

## Urban Network Analysis

**Modeling Paradigm 2**

**Graph data (nodes & edges) – Network representations – spatial graphs – centrality metrics – power-law distributions – applications to mobility & spatial networks**

## Machine Learning in Cities

**Modeling Paradigm 3**

**GLIN data – supervised learning – unsupervised learning – deep learning (images & graph)**

# Multiple data structures in cities

| | |
|---|---|
| Graph (road, building) | Lang (Yelp, Twitter) |
| Image (Street and satellite) | Num (Travel survey) |

Four basic data structures: GLIN



- Uniqueness of cities: connecting GLIN
- Uniqueness of deep learning to unify GLIN (GNN, RNN, CNN, ANN)

# Logic in data structures vs. modeling paradigms

More complex modeling paradigms

More complex data structures (GLIN)

| | 1970s - Urban Statistical Analysis | 1990s - Urban Network Analysis | 2010s - Machine Learning in Cities |
|---|---|---|---|
| Numbers | ✓ | ✓ | ✓ |
| Graphs | × | ✓ | ✓ |
| Images | × | × | ✓ |
| Languages | × | × | ✓ |

# Goals of this course

### 1. Learn analytical perspectives

- Primary modeling paradigms
- Similarities & differences
- NOT: in-depth math & single analytical approach

### 2. Learn urban applications

- Mobility & economic applications.
- Generally applicable to others
- NOT: single topic focus

### 3. Learn computational practice

- Python for urban analytics & applications
- Take some efforts!
- NOT: algorithm class

**Values**

- **To Master Students**. (1) General analytical perspectives, (2) broad urban applications, and (3) computational practices.
- **To PhD Students**. Not in-depth into single approach & topic, but the broad roadmap helps to identify your positioning.

Q: What does a chef need to learn for cooking?

# How to think about the knowledge components in this course?

- Analytical Perspectives
  Recipe
- Urban Applications
  Ingredients
- Computational Practice
  Cookers

# Some uncertainty in the field of urban analytics

- People use different names. (e.g. urban computing, urban AI, urban data science, etc.)
- People discuss different contents. (e.g. GIS, agent-based simulation, etc.)
- It is simply because this field is new.

# Part 3. How will this course be taught?

# General Course Information

- Course: Introduction to Urban Analytics
- Course number: URP6931
- Time: Tuesday 1:55-4:55pm (1:55-2:45pm; 3:00-3:50pm; 4:05-4:55pm)
- Location: RNK 0210
- Instructor: Shenhao Wang
- Office hours: Wednesday 2-3pm (Starting Jan 18)
- Email: shenhaowang@ufl.edu (expected response time: about 48 hours)

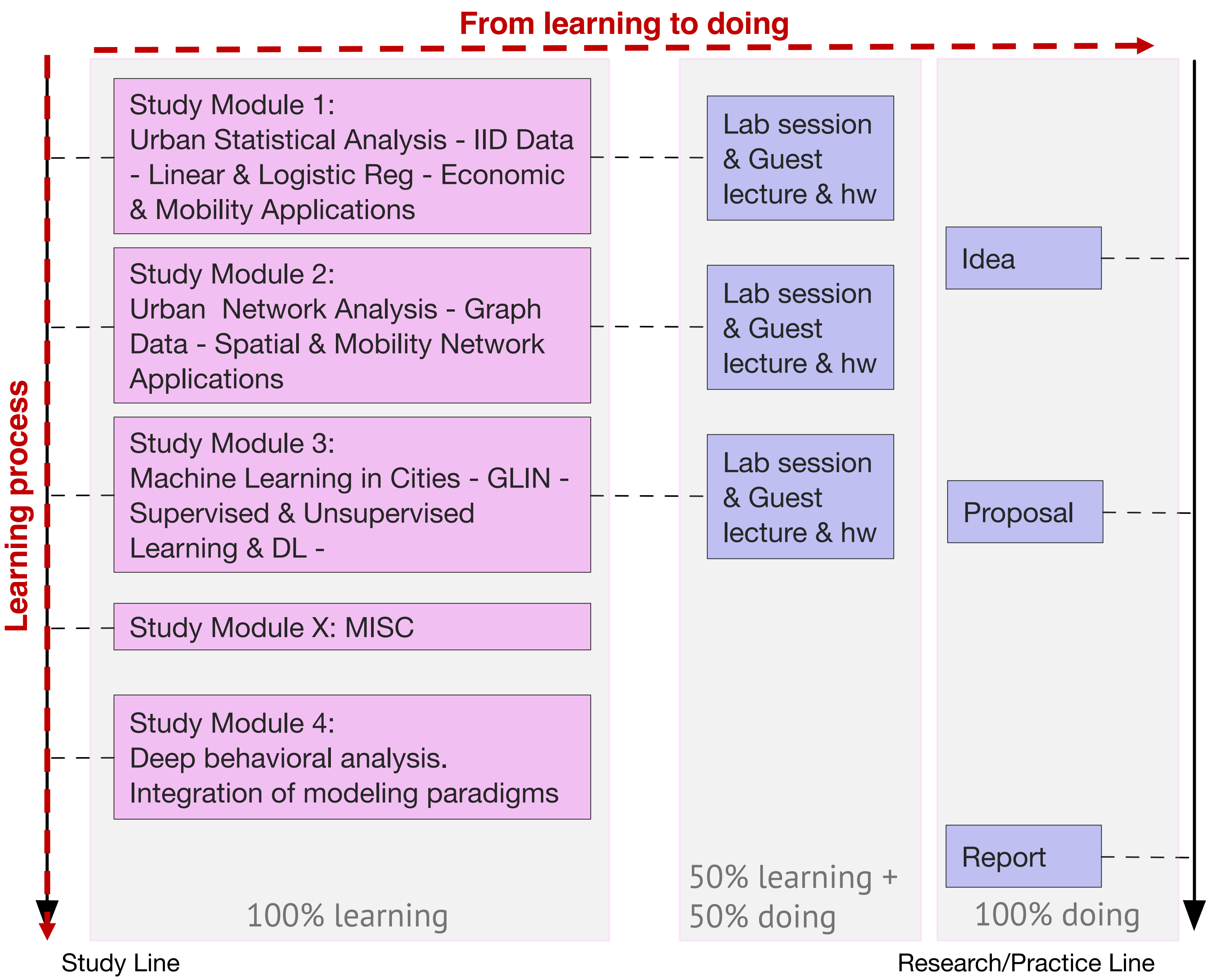
# Logic of designing the class

**1. Learning process**

- Individual modeling paradigms: study modules 1-3
- Integration of model paradigm: study module 4

**2. Learning + Doing**

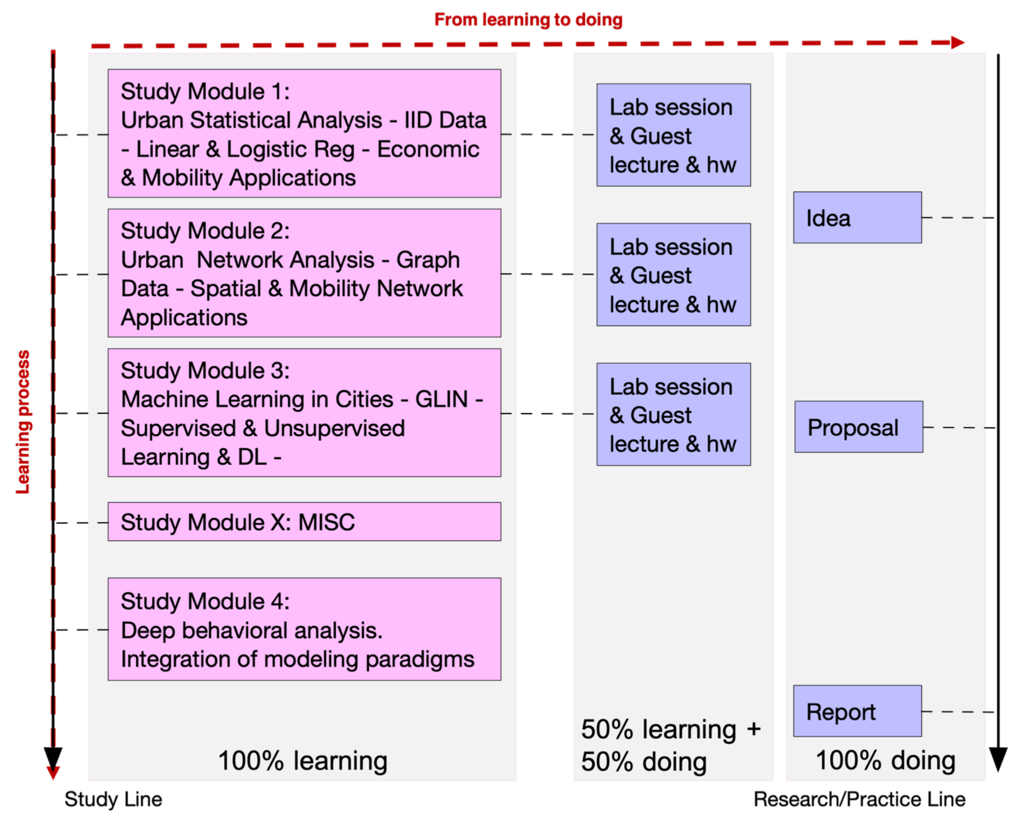
- Lectures. 100% learning
- Lab + Guest lecture + Psets: 50% learning + 50% doing
- Project: 100% doing




Framework of Introduction to Urban Analytics

# Tentative course schedule

Work in progress




Framework of Introduction to Urban Analytics

| Week | Dates | | Guest lectures | Lab sessions | Psets | Project |
|---|---|---|---|---|---|---|
| 1 | Jan 10 | class overview | | | | |
| 2 | Jan 17 | review: python and statistics | | Y | | |
| **Module 1: Urban statistical analysis** | | | | | | |
| 3 | Jan 24 | linear regression & economic application | | Y | | |
| 4 | Jan 31 | logistic regression & travel choice analysis | | Y | | |
| 5 | Feb 7 | regression critiques in urban analytics | Y | | hw 1 out | Idea due |
| **Module 2: Urban network analysis** | | | | | | |
| 6 | Feb 14 | urban networks and centrality metrics | | Y | | |
| 7 | Feb 21 | network methods: scaling, community, and spatial regression | Y | | hw 2 out, hw1 due | |
| **Module 3: Machine learning in cities** | | | | | | |
| 8 | Feb 28 | supervised learning: classifications and regression revisited | | Y | | |
| 9 | Mar 7 | unsupervised learning: clustering | | | hw 3 out, hw2 due | |
| NA | Mar 14 | no class (Spring Break) | | | | |
| 10 | Mar 21 | deep learning basics | | Y | | Proposal due |
| 11 | Mar 28 | deep learning with images and graphs | Y | | hw 3 due | |
| **Module X: MISC** | | | | | | |
| 12 | Apr 4 | optimization, simulation, and justice | | | | |
| **Module 4: Behavioral analysis – integration of three paradigms** | | | | | | |
| 13 | Apr 11 | deep choice models | | | | |
| 14 | Apr 18 | deep choice models | | | | |
| 15 | Apr 25 | final presentation & course evaluation | | | | Report due |

# Tentative course schedule

Limitations (IMPORTANT)

Module 1: Six courses (Shenhao)
Module 2: Three courses (Shenhao)
Module 3: Eight courses (Shenhao)

**Gateway Class**

- Open doors for urban analytics
- But not in-depth.

| Week | Dates | | Guest lectures | Lab sessions | Psets | Project |
|---|---|---|---|---|---|---|
| 1 | Jan 10 | class overview | | | | |
| 2 | Jan 17 | review: python and statistics | | Y | | |
| **Module 1: Urban statistical analysis** | | | | | | |
| 3 | Jan 24 | linear regression & economic application | | Y | | |
| 4 | Jan 31 | logistic regression & travel choice analysis | | Y | | |
| 5 | Feb 7 | regression critiques in urban analytics | Y | | hw 1 out | Idea due |
| **Module 2: Urban network analysis** | | | | | | |
| 6 | Feb 14 | urban networks and centrality metrics | | Y | | |
| 7 | Feb 21 | network methods: scaling, community, and spatial regression | Y | | hw 2 out, hw1 due | |
| **Module 3: Machine learning in cities** | | | | | | |
| 8 | Feb 28 | supervised learning: classifications and regression revisited | | Y | | |
| 9 | Mar 7 | unsupervised learning: clustering | | | hw 3 out, hw2 due | |
| NA | Mar 14 | no class (Spring Break) | | | | |
| 10 | Mar 21 | deep learning basics | | Y | | Proposal due |
| 11 | Mar 28 | deep learning with images and graphs | Y | | hw 3 due | |
| **Module X: MISC** | | | | | | |
| 12 | Apr 4 | optimization, simulation, and justice | | | | |
| **Module 4: Behavioral analysis – integration of three paradigms** | | | | | | |
| 13 | Apr 11 | deep choice models | | | | |
| 14 | Apr 18 | deep choice models | | | | |
| 15 | Apr 25 | final presentation & course evaluation | | | | Report due |

# Example. Feb 28 – Supervised Learning

1. Analytics

Lecture: Supervised learning

2. Application

Lecture: Travel demand prediction

3. Computation

Lab: Python scikit-learn with census data

Homework 3

1 + 2 + 3

(different data set)

# Grading

| Components | Percentage of final grade |
|---|---|
| Course participation | 10% |
| Problem sets (3) | 45% |
| Project | 45% |

1. **Course participation (10 pts)**: Attend the class. Ask for approval of absence.
2. **Problem sets (45 pts)**: contents from the lecture + scripts from the lab + different data set

- Pset 1 (15 pts) – Regression
- Pset 2 (15 pts) – Network analysis
- Pset 3 (15 pts) – Machine learning

3. **Project (45 pts)**: team project (2~3 students per team) with free choice in topic but should focus on one modeling paradigm taught in this class.

- Idea (5 pt). Limit to 1 page.
- Proposal (10 pt). Limit to 3 pages.
- Final paper (30 pt). Limit to 8 pages.

# Prerequisites: math and coding

- **No strict prerequisite**
  - **Motivation** - I want to support planning students, who typically have relatively limited math training
  - Not sure if this is the best…

- However, it is **highly recommended** that you used Python before, and learnt basic probability, statistics, and linear algebra.
  - Took an **Introduction to Python** course
  - Used Python packages, e.g., numpy, pandas.
  - Math: undergraduate probability and statistics

- If not, I will help you review the critical background knowledge, but the learning curve will be **steep**.

# Acknowledgement

1. First time to teach such a class.
2. Uncertainty: opportunities vs. risks
   - Syllabus
   - Guest lectures
   - Lab sessions and Psets
   - Project requirements
3. Your feedbacks are really important.

# Take-aways

### 1. Why?

1. Challenges and opportunities
2. Formalize and generalize planning tradition and intuition
3. Trends in department, university, and society

### 2. What?

1. Three modeling paradigms
2. Data representations
3. Models + Application + Computation

### 3. How?

1. Four study modules + X
2. From learning to doing
3. Gateway class

# Q&A