Big Data for Cities
Week 12

Curt Savoie
Connor McKay

Agenda

- **Recap of last week**
- **Advanced Topics**
  - **Data Mining Basics**
  - **Industry Survey**
- **R Demo**

Recap

- **Feeling good about the final scope / questions?**

Data Mining / Machine Learning

- ## What is data mining?
  - **Data Mining** is an analytic process designed to explore **data** (usually large amounts of **data** - typically business or market related - also known as "big **data**") in search of consistent patterns and/or systematic relationships between variables, and then to validate the findings
- ## What is machine learning?
  - is a field of computer science that gives computers the ability to learn without being explicitly programmed

Data Mining / Machine Learning

- A few major areas of algorithms
  - Association Rules
    - What things go together; market basket analysis
  - Classification
    - Tagging in a sense; spam email for example
  - Clustering
    - Creating groupings

Data Mining / Machine Learning

- Interesting Machine Learning Techniques
  - Neural Nets
  - Deep Learning

Industry Stuff

- Let's review what you've learned this semester first
    - R
    - Useful statistics techniques
    - Urban data and city operations
    - Visualization
    - Scoping of project plans

Industry Stuff

- Government
  - Cities, states, federal
- Real Estate / Development
- Research / Academics
- Advocacy
- Urban/Regional Planning
- Finance / Economics
- Public Health
- Transportation

For Next Week

- **Guest lecture next week!**
- **Reading on theory and practice**
  - nothing
- In R
  - Work on final projects