Experience with an Interdisciplinary Competition-based Cybertraining Workshop

Bing Wang, Suining He, Chuanrong Zhang, Abdul-Wassay Queeshi, Weidong Li, Sanguthevar Rajasekaran, **Wei Wei,** Elizabeth Howard

School of Computing
Department of Geography
School of Education

UCONN | UNIVERSITY OF CONNECTICUT

Introduction to Cyberinfrastructure (CI)

- CI includes computing systems, data, software, visualization, and people
- Supports data-driven research and scientific discovery
  - e.g., Satellite imagery, IoT sensors, GPS data from smartphones
- Challenge
  - Education and workforce development lag behind CI's importance

Workshop Motivation

- Growing need for CI skills in research and industry
- Limited training in handling large-scale spatio-temporal data
- Goal
  - Train students to use CI for research and develop innovative applications
- Approach
  - Interdisciplinary, competition-based workshop to foster skills and collaboration

Workshop Overview

- 2 weeks, Spring 2024
- Participants: 10 students
  - 5 undergrad, 5 grad
  - from CSE and Geography
- 4 interdisciplinary teams
  - 2–4 students each
- Theme
  - Efficient management of bike-sharing systems
  - Using NYC Citi Bike data

Interdisciplinary Design

- Why Interdisciplinary?
  - CI projects require diverse expertise
- CSE Students
  - Skilled in computing, machine learning
  - Limited spatio-temporal data experience
- Geography Students
  - Proficient in spatial analysis
  - Limited computing infrastructure knowledge
- Goal: Foster collaboration to leverage complementary skills

Competition-based Format

- Why Competition?
  - Stimulates interest, enhances learning (Burguillo, 2010)
- Benefits
  - Encourages critical thinking, teamwork, and innovation
- Mitigating Negatives
  - Team-based competition reduces stress, focuses on collaboration
- Judging Criteria
  - Technical merit (40%)
  - Team collaboration (40%)
  - Presentation quality (20%)

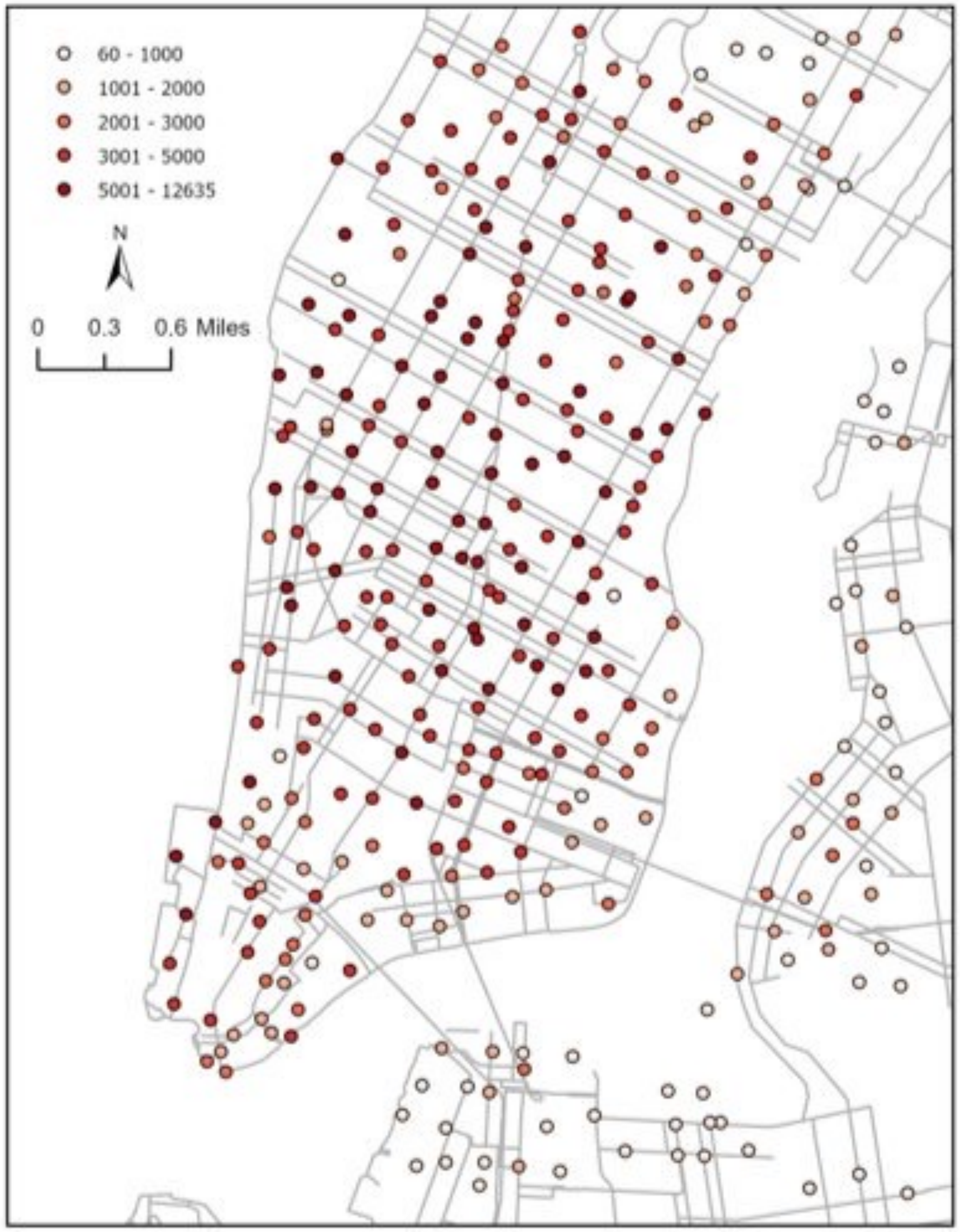
Workshop Problem: Bike-Sharing Systems

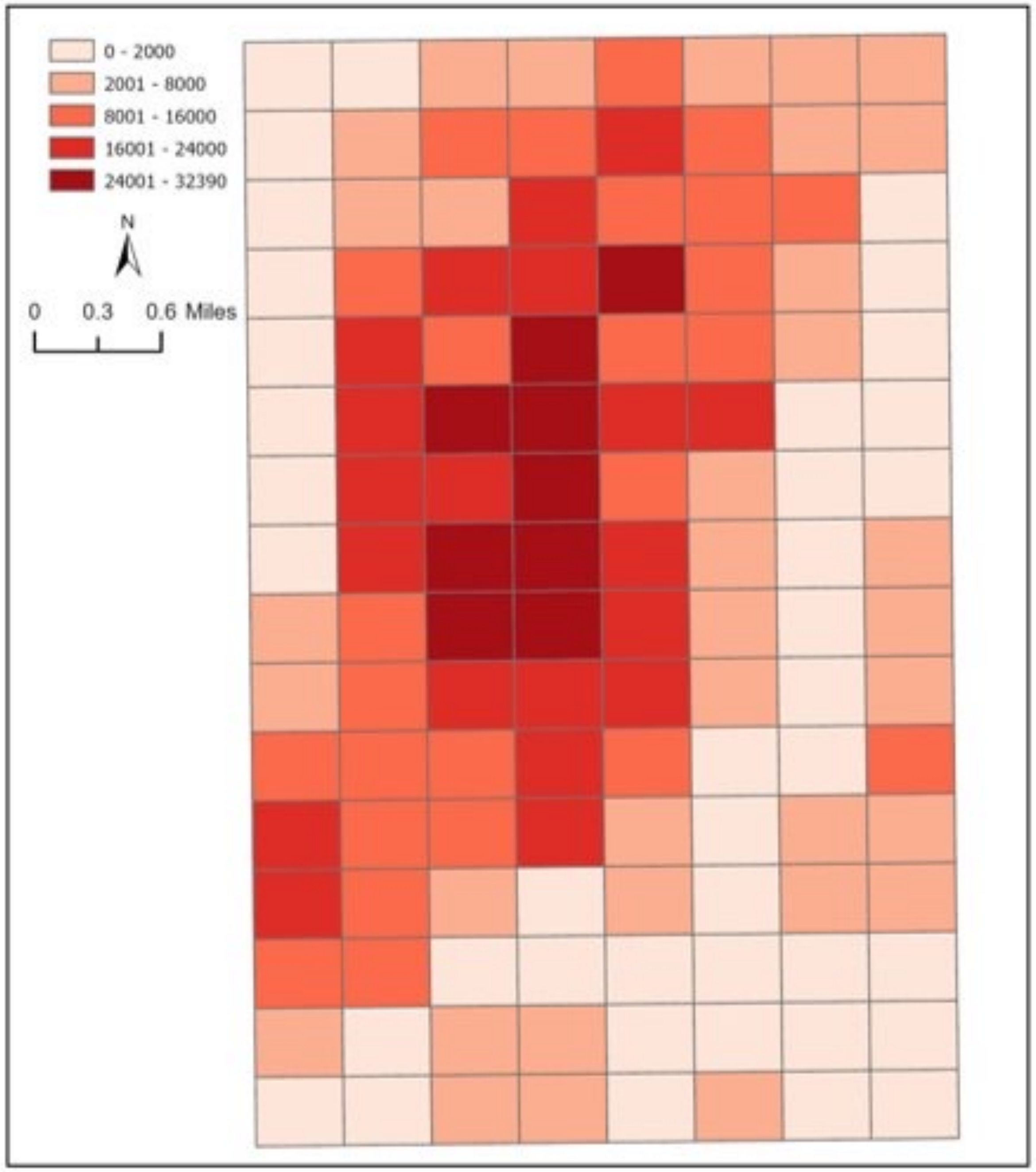
- Predict bike flow (pick-ups/returns) for efficient bike-sharing management
  - Supports urban mobility, reduces rebalancing costs
- Challenges
  - Spatial and temporal variations, influenced by urban layout, weather, etc.
- Task
  - Analyze spatio-temporal data, develop predictive models

Dataset Description

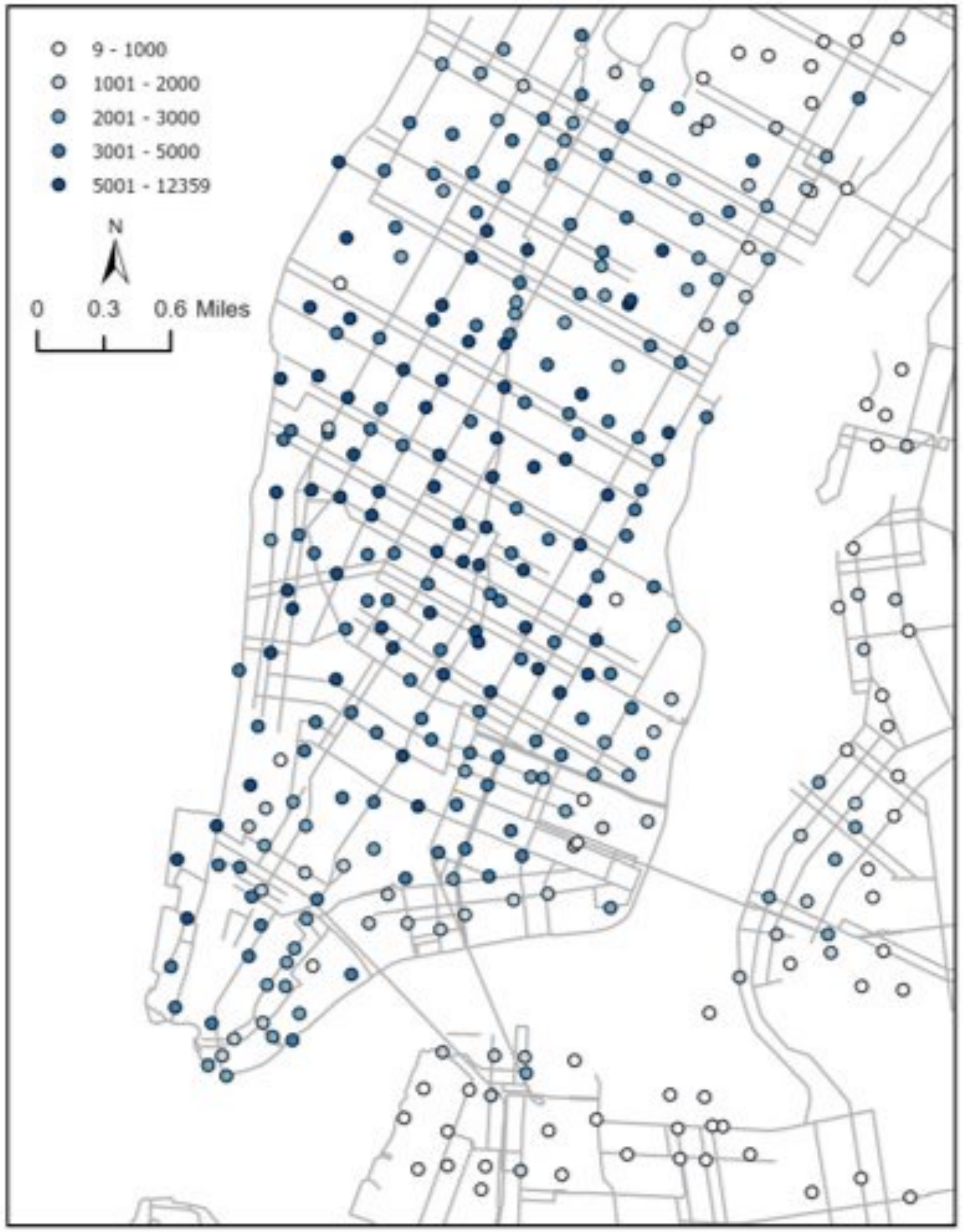
- Source: NYC Citi Bike Dataset (Oct 2019)
- Data Types
  - Raw data (.csv): Bike stations, trips, rider info
  - Processed data (.h5): Spatio-temporal tensor (16×8 grid, 30-min intervals)
- Additional Data
  - Teams encouraged to use external datasets (e.g., crime rates, bike lanes)

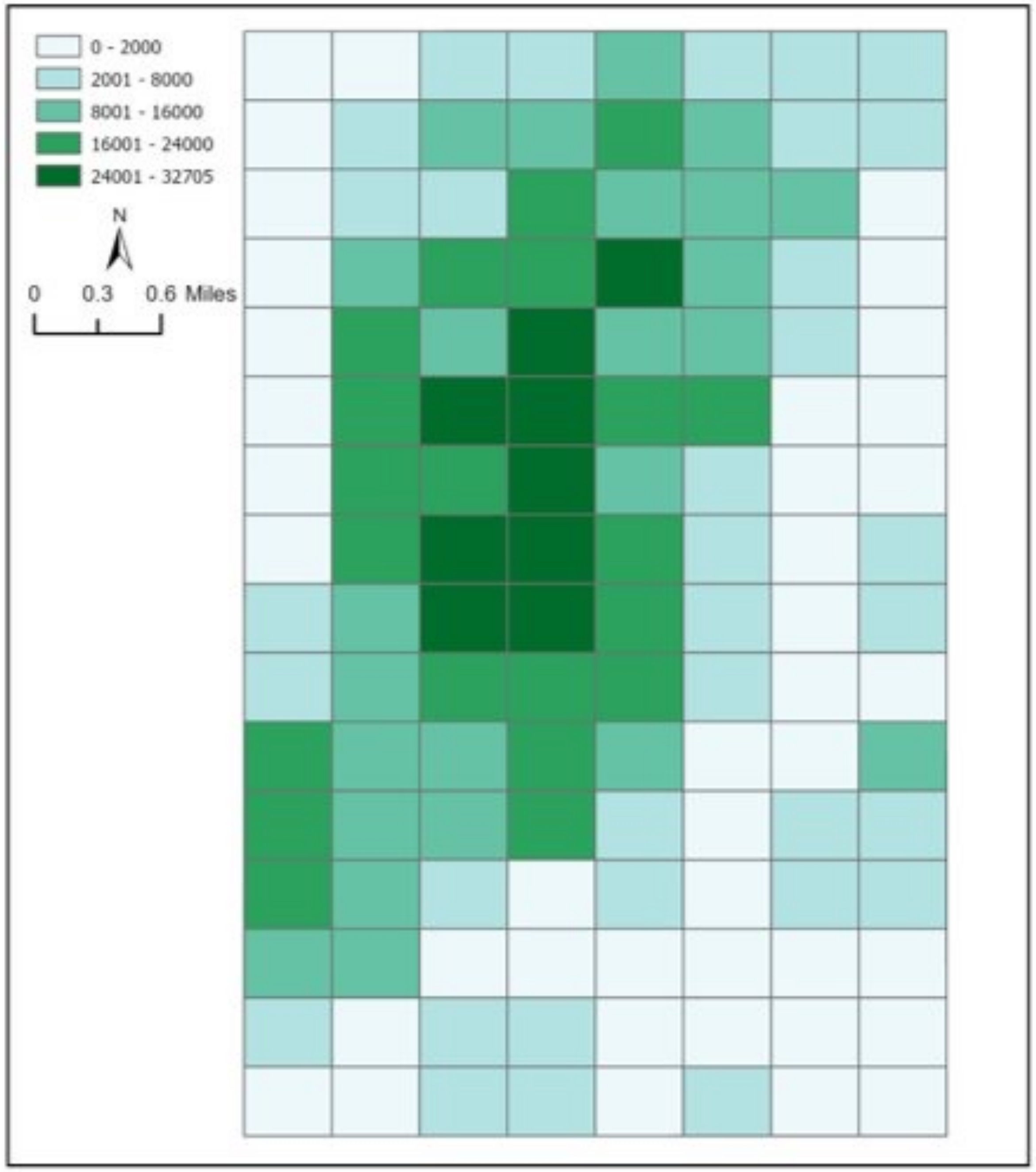
Data: # of pick-ups by start stations and grid cells





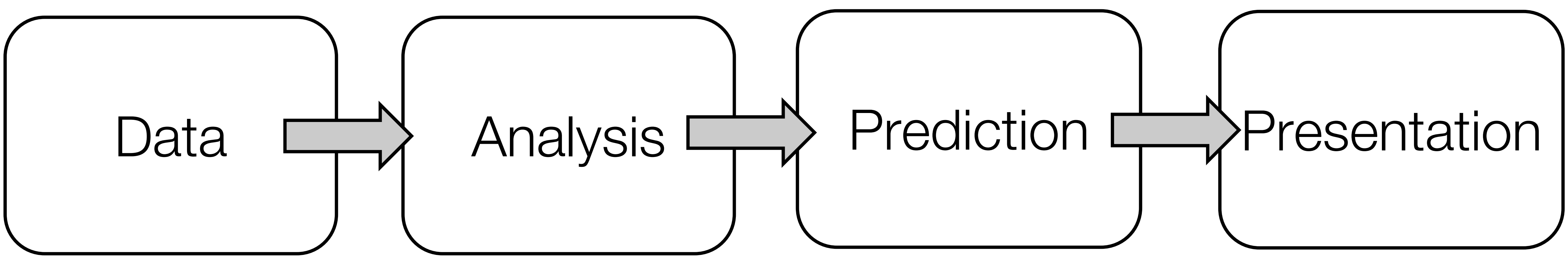
Data: # of returns by stop stations and grid cells





Workshop Activities

- Tasks
  - Find/download relevant data
  - Perform spatio-temporal analysis (e.g., GIS visualizations)
  - Develop machine learning models for bike flow prediction
- Process
  - Teams collaborated on open-ended research, presented results



Training and Support

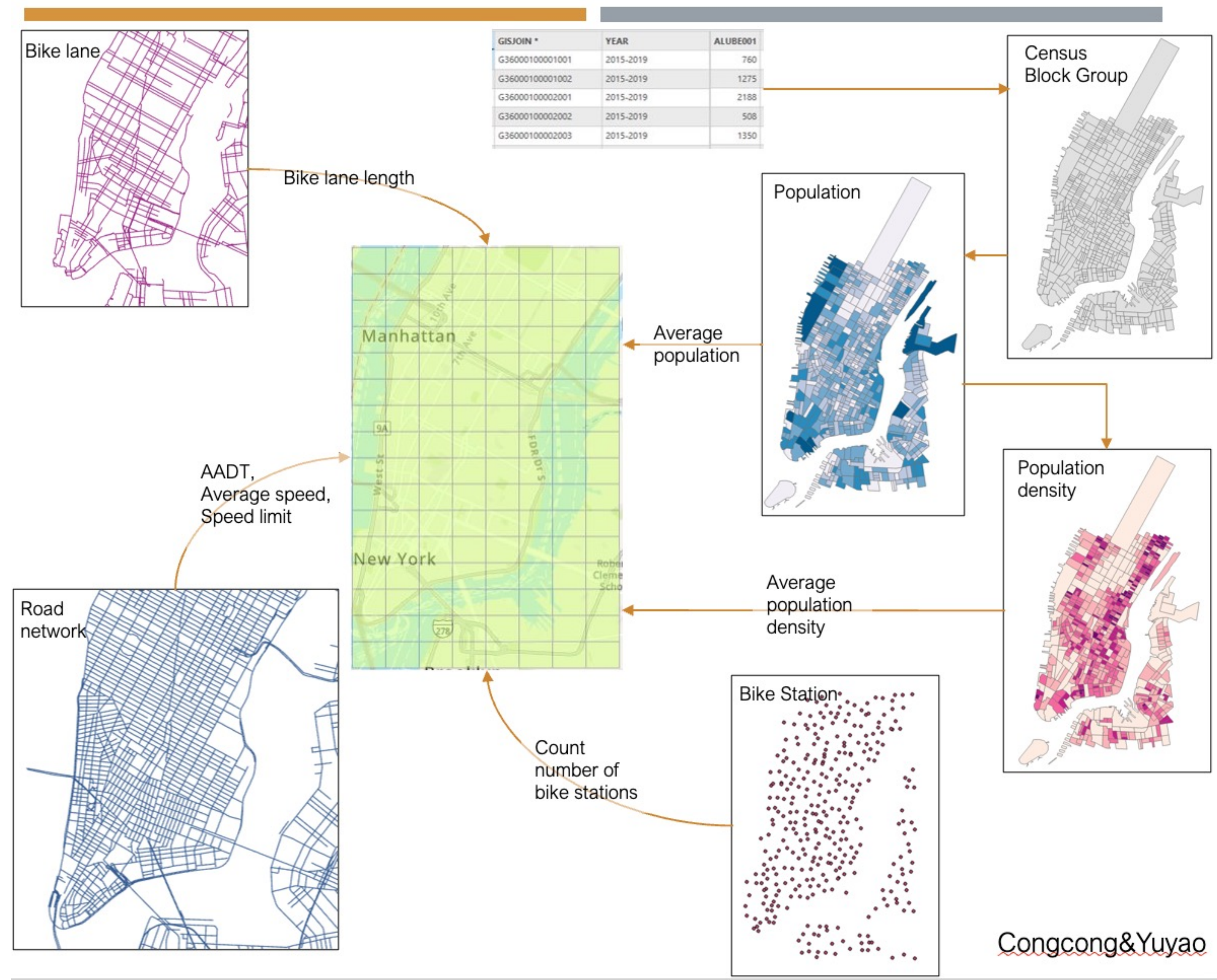
- Materials Provided
  - Reference papers (CSE and Geography)
  - Tutorial on data analysis
  - Guidelines for interdisciplinary collaboration
- Coaching
  - Two sessions per team (week 1: planning, week 2: feedback)
- Purpose
  - Support students in research and teamwork

Spatio-Temporal Analysis

- Tools
  - ArcGIS for visualization (heat maps, flow maps, space-time cubes)
- Analyses
  - Spatial: Identified high-demand areas, popular routes
  - Temporal: Analyzed usage patterns (daily, weekly)
  - Network: Evaluated connectivity, integration with public transit
- Findings
  - Insights on user demographics, accessibility gaps

Spatial Data From Student Team

- Bike Station (raw data)
  - Count within each cell
- Bike lane (NYDOT)
  - Total length within each cell
- Traffic volume (NYDOT)
  - AADT, average speed, and speed limit
  - Mean value within each cell
- Population (NHGIS)
  - Population
  - Derived population density
  - Mean value within each cell

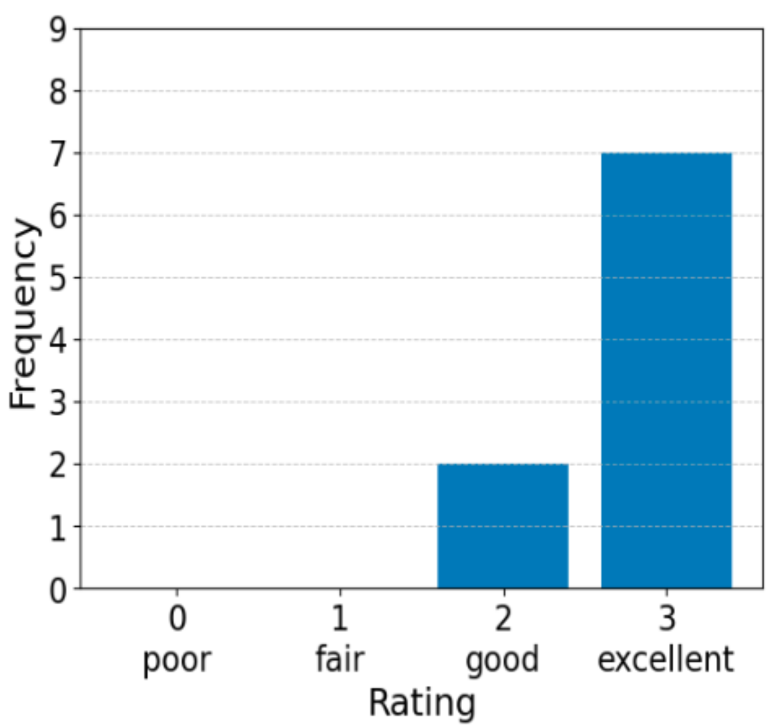


Machine Learning Predictions

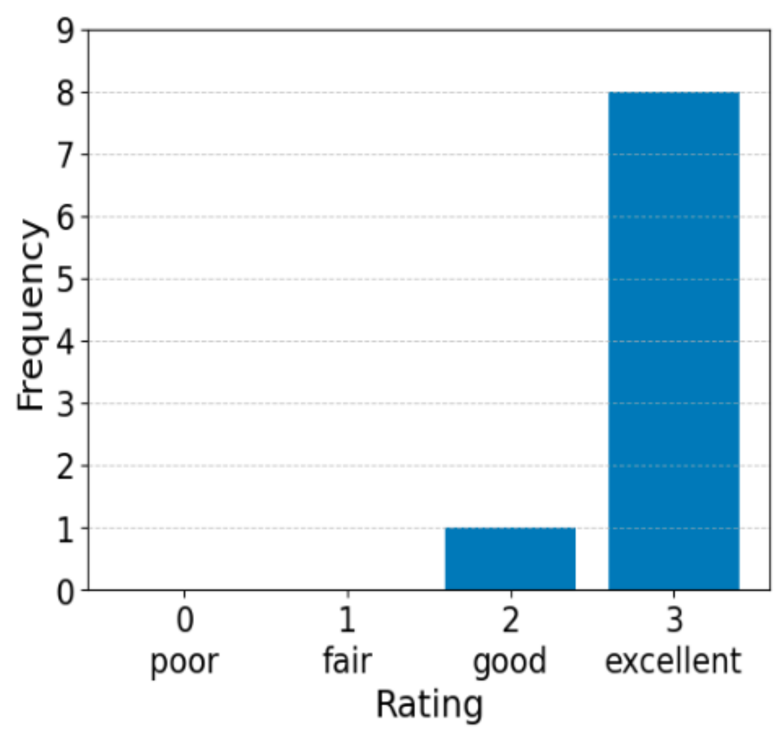
- Approach
  - Teams developed models (LSTM, GRU, dense layers)
  - Ablation studies to assess feature impacts
- Features
  - Incorporated population density, weather, bike lane data, ...
- Evaluation Metrics
  - Mean Square Error (MSE) on four weeks of test data

Survey Results (Post-Workshop)

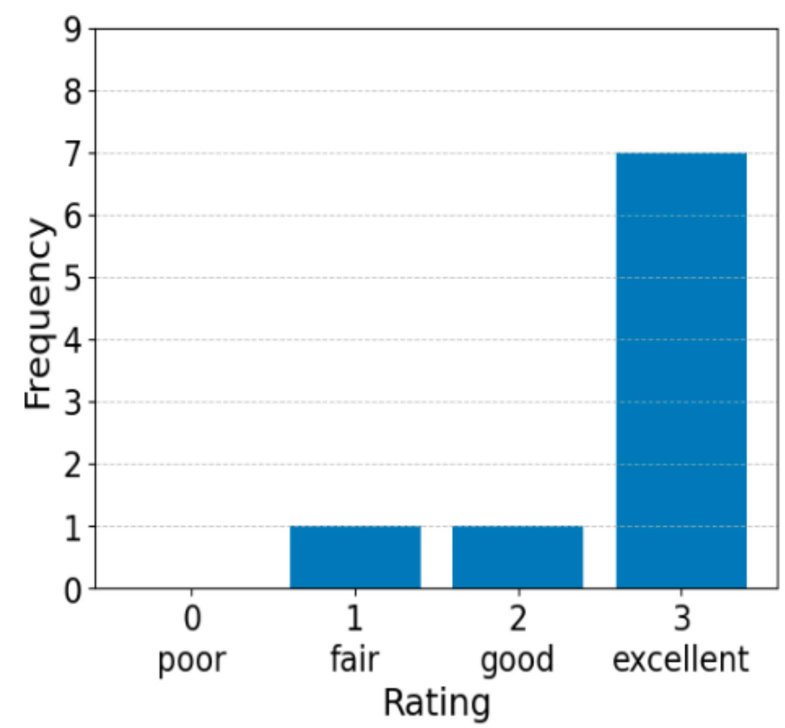
- Overall Experience: 7/9 rated “excellent,” 2/9 “good”
- Content: 8/9 “excellent,” 1/9 “good”
- Teamwork: Mixed (2 fair/good, 7 excellent)



(a) Overall experience.



(b) Content.

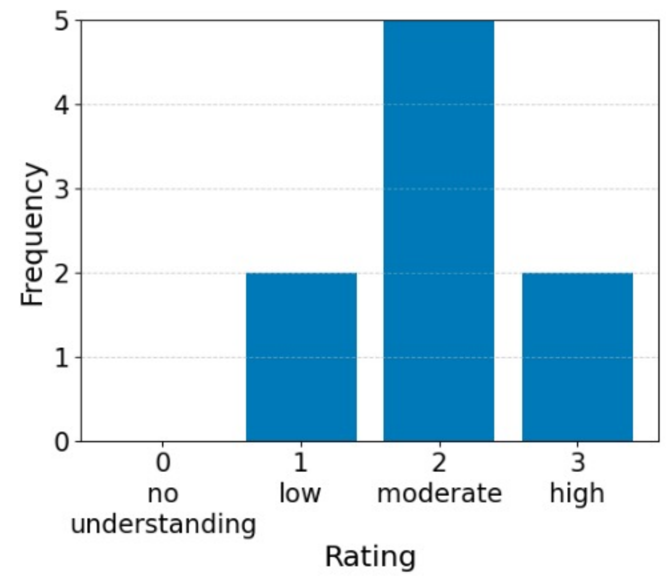


(c) Teamwork experience.

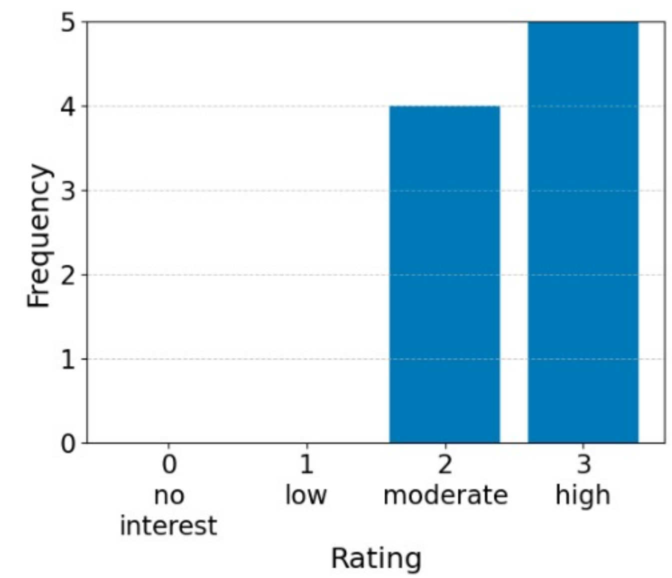
Survey Results (Pre vs Post Workshop)

- Improvements:
  - CI Understanding:
    - All moderate/high, shifted higher
  - Interdisciplinary Interest:
    - All reported “high”

Pre-workshop

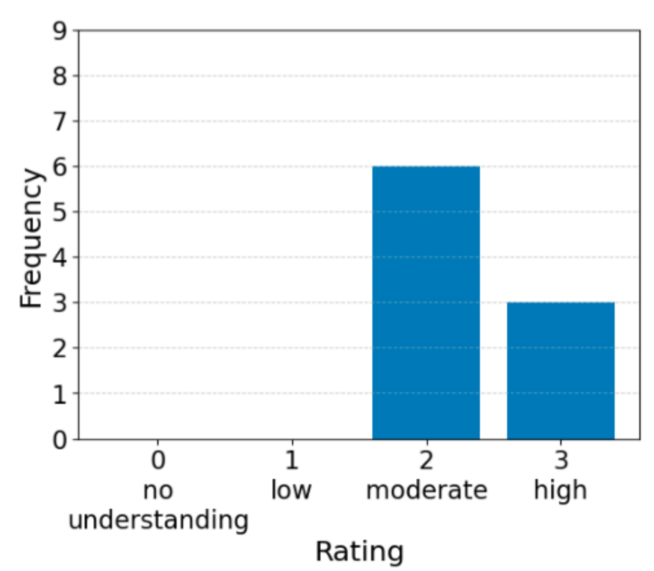


(a) Understanding of CI.

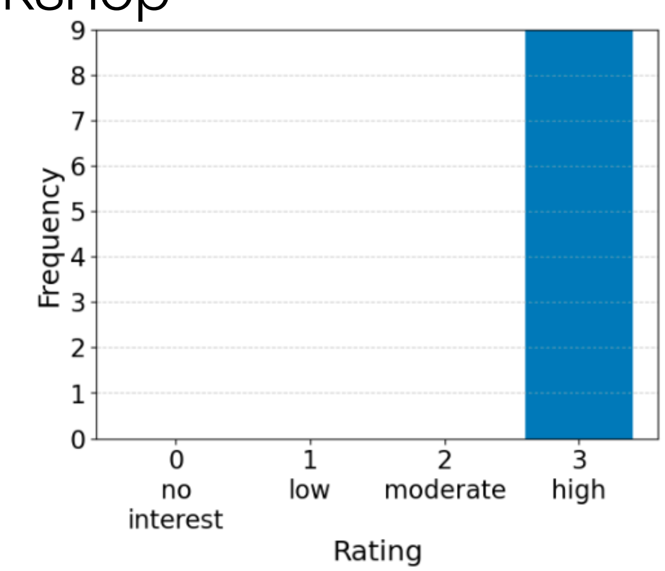


(b) Interests in interdisciplinary work.

Post-workshop



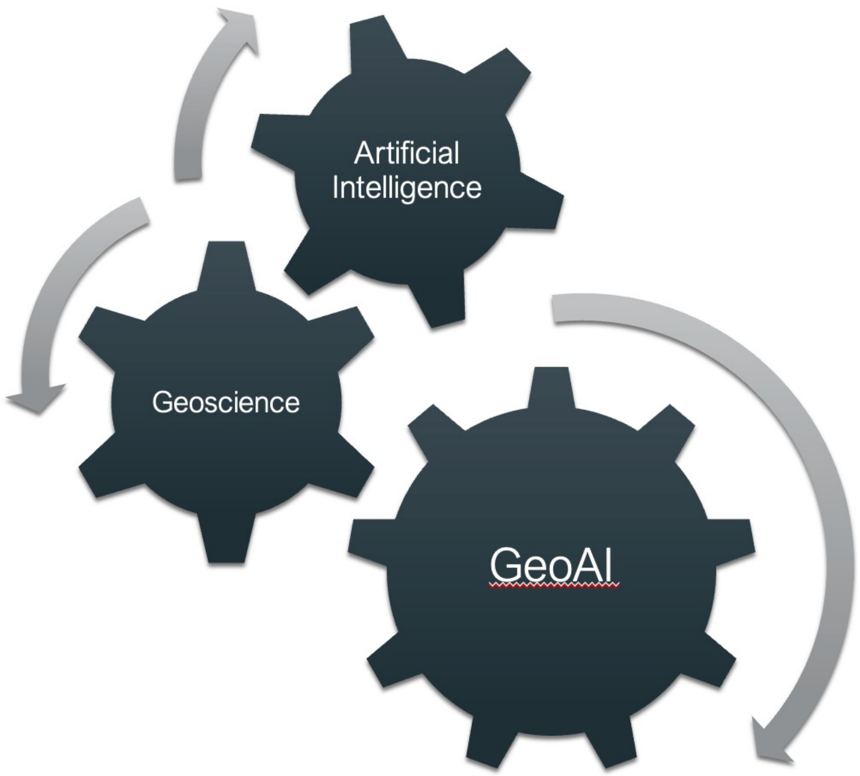
(a) Understanding of CI.



(b) Interests in interdisciplinary work.

Team Project Highlights

- Visualizations
  - 2D/3D maps, web-based dashboards
- Innovations
  - Crime rate clustering, multi-scale analysis
- Findings:
  - Primary user age groups, trip characteristics
  - Unsafe station locations, bike lane impacts
- Interdisciplinary Insight
  - GeoAI benefits (AI + spatial context)



Interdisciplinary Collaboration

- Participants valued learning from diverse teammates
- In-person meetings enhanced idea exchange
- Increased interest in interdisciplinary work (survey results)
- Found limited time for team coordination

Lessons Learned

- Successes
  - Competition format engaged students
  - Interdisciplinary teams fostered learning
  - Undergraduates excelled in research tasks
- Challenges
  - Short duration limited exploration
  - Late team formation hindered collaboration

Recommendations

- Extend Duration
  - 3–4 weeks for deeper research
- Earlier Team Formation
  - Allow more time for bonding
- Sustained Collaboration
  - Link to independent studies, REU programs
- Scale Up
  - Validate with larger cohorts, explore new teaming strategies

Future Directions

- Long-Term Impact
  - Encourage post-workshop projects (e.g., publications)
- Broader Reach
  - Expand to more students, disciplines
- Research for Undergrads
  - Integrate into curricula or research programs
- Evaluation
  - Conduct more workshops to refine approach

Conclusion

- Organized workshop that successfully trained students in CI and interdisciplinary collaboration
- Exceeded expectations with innovative projects
- Participants provided positive feedback
- Takeaways
  - Competition-based format is effective for CI education
  - Interdisciplinary collaboration stimulated interests in future interdisciplinary work
- Apply lessons to future workshops for broader impact

Acknowledgment: supported by NSF OAC-2118102