Development, Deployment, and Assessment of a New Educational Paradigm for Transportation Professionals and University Students

A Collaboration of the Region X Transportation Consortium

Understanding and Communicating Multimodal Transportation Data

Portland State University
Christopher Monsere

Course Design Document

1. INTRODUCTION

This course will introduce students to appropriate research methods for using transportation data sets and communicating the results of their work to a broad audience. The course content includes:

(a) selections of the appropriate graphical method
    a. knowledge-based decisions on selections for best perceptions
(b) managing, extracting, and filtering large-scale data
(c) understanding types and dimensions of data (time resolution, discrete, continuous, and aggregations)
(d) techniques for visualizing data and exploratory analysis
(e) basic statistical analysis applied to transportation problems (public transportation, traffic, safety, freight, bicycle performance) using open-source script-based statistical tools (R) and databases (PostgresSQL)
(f) selection of appropriate analysis technique
(g) presentation of material in a technical summary

This is a gateway course; the knowledge gained in this course will be applied throughout the remaining graduate curriculum.

2. AUDIENCE ANALYSIS

Students taking this course will have had an introductory transportation course, an undergraduate course in statistics and probability, and an engineering problem solving course with an exposure to programming logic. Three audiences are envisioned for this course:

(a) Graduate-level civil engineering students with an emphasis in transportation, in their first quarter.
(b) Transportation professionals with a desire to expand their knowledge of data analysis
(c) Advanced senior undergraduate civil engineering students with necessary skills and permission of the instructor.

The course uses the open source language R. Use of the PostgreSQL database will require comfort with various computing platforms (Unix, Windows) including the installation of software, downloading and installing web-based technologies.

The long-term behaviors, roles, and way of being will be supported by this course:

(a) Problem solver
(b) Researcher

(c) Communicator
(d) Collaborator

3. COURSE VISION

Preparing future transportation professionals to interact, explore, analyze, and explain multimodal transportation data. Students who complete this course will be comfortable with data.

4. COURSE ENVIRONMENT

This course will be offered in the fall quarter to capture students in their first quarter of graduate work. This will overlap with the semester system. The course is primarily designed to be taught with a local instructor or facilitator, with the option to include distance components if desired.

Credit Hours	***Student Time Commitment***	***Types of Student Work***
In the quarter system, a 4-credit course. This is a 10 week course that meets for 1:50 minutes twice a week.	*Approximately 36 contact hours in class with approximately 15 assigned hours out of class effort.*	• *Individual in-class and homework* • *Group in-class and homework* • *Computer lab work*

5. LEARNING OUTCOMES

Table 1 Learning Outcomes

1. Competencies	***2. Movement***	***3. Experience***	***4. Integrated performance***
1. *Identify ethical issues likely to be encountered in research and data* 2. *Identify transportation research organizations and*	1. *Formulating inquiry questions to understand and describe the data* 2. *Increase programming proficiency in R including skills*	1. *Receiving feedback on a group project from peer reviewers.* 2. *Communicating ideas and knowledge via various*	1. *Use data analysis, hypothesis testing, and visualization tools to answer and explain complex transportation problems.*

explain the research they sponsor *3. Explain and identify the basic types of data (continuous, discrete, nominal, ordinal) and variable types (independent, dependent, categorical)* *a. What does this mean for analysis* *b. How to present these data* *c. Be able to select appropriate tool* *4. Construct simple and complex SQL queries to extract data from large scale RDMS* *a. Data filtering process* *b. SQL* *c. Filling in missing data* *d. Preparing the data before analysis* *5. Make judgments about the correlation between one or more variables (using linear regression)* *6. Use exploratory techniques to identify outliers and "bad" data*	*and logic to develop and implement scripts and algorithms* *3. Apply visualization techniques to gain understanding of large-scale data.* *4. Select and conduct appropriate statistical tests to test hypothesis at different levels of significance.* *a. Hypothesis testing (define scope of type of tests)* *b. Understanding the assumptions of various statistical tests* *c. Distribution fitting* *5. Prepare a summary of analysis (including graphics) for technical audience.*	*mediums (in-person spoken, written, web-based).* *3. Active exploration of interacting with transportation data archives to answer authentic problems.*	

6. KNOWLEDGE TABLE

Table 2 Knowledge Table

Concepts	***Process***	***Tools***	***Context***	***Ways of Being***
1. Quantitative and qualitative data *2. See statistical inventory of concepts* *3. Visualization techniques and graphical presentation (Tufte's guidelines)* *a. Different graph types* *b. Use of color* *c. R* *d. Avoiding bias* *e. Preparation of graphs for appropriate audience* *4. Data formats*	*1. Research methods* *2. Data Analysis* *3. Hypothesis testing* *4. Visualization* *5. Communication Theory and Process* *6. Debugging and syntax*	*1. Relational Database (PostgreSQL)* *2. Excel spreadsheet* *3. Unix shell* *4. R language and environment for statistical computing and graphics* *5. Presentation Software (PPT, Webinar)*	*1. Different datasets, biases, etc. (what are the different elements, Simple datasets, Larger complex multivariate datasets)* *2. Application – transportation (bus, traffic, safety, wim, ped, bicycle, probe)* *3. Discussion forums, face to face and online* *4. Collaborative grading (peer review) and feedback*	*1. Researcher* *2. Analyst* *3. Communicator* *4. Collaborator*

7. LEARNING SKILLS

Learning skills are discrete entities that are not knowledge or discipline specific and are therefore transferable across many areas of human experience. They can be consciously improved and refined. Once they are improved, the rate and effectiveness of overall learning increases. No matter what a person's age or experience, learning skills have no upper bound. They can always be improved. Learning skill development is usually triggered by a learning challenge of some kind that forces the student to use or further develop a certain set of skills and is facilitated through mentoring. Candidate skills can come from three domains: cognitive, social, and affective.

Cognitive		***Social Skill***		***Affective Skill***	
Filtering – selecting data based on criteria		Formatting a message – selecting a mode or style that fits the purpose		Being curious – wanting to find out more Being open– welcoming, and expecting to find novelty	
Recognizing patterns – perceiving consistent repetitive occurrences		Illustrating – enhancing a message with images, tables or drawings		Persisting – continuing despite difficulties	
Testing perceptions – verifying based on interpretations				Being self-disciplined – persisting regardless of emotions	
Validating sources – verifying based on credibility				Aligning with social values – acting according to mutually empowering ethics	
Controlling errors – verifying based on procedures					
Identifying inconsistency – detecting outliers/anomalies					
Identifying assumptions – examining preconceptions/biases					
Recognizing the problem – stating what is wrong or missing					
Selecting tools – finding methods to facilitate solution					
Stating research questions –					

asking empirically answerable questions					
Selecting methods – determining research procedures					
Testing hypotheses – discerning significant effects					

8. ASSESSMENT AND EVALUATION TOOLS

Activities are also shown as in class or out of class. The following structure will be used to assess activities:

1. Participation Activities
    a. These activities require quick assessment and feedback. You will receive credit for completing these activities.
2. Annotated Code Activities
    a. In these activities you be asked to only submit a script or code file that contains comments and demonstrates active exploration of the objectives within the activity.
3. Peer Assessment Activities
    a. Some activities will require you to assess the work of your fellow students. In these activities, your performance will be based on your work assessed by the instructor, your feedback to peers, and your peers' assessment of your work.
4. Short Response Activities
    a. Many activities are structured such that you respond to a set of questions. You will receive credit both for completing these activities and for the depth and detail of your responses. We will attempt electronic submittal and feedback for these activities.
5. Discovery Activities
    a. These activities require you to build on knowledge and skills introduced to you in previous activities. These activities will be open ended and you will receive credit for completing these activities and for the creativity of your exploration.

There will also be a final project which is an independent structured analysis which you will select from a set of open-ended questions devised by the instructor presented in Chapter 6. This project will serve as the final assessment that the student has made progress in developing knowledge and skills in this class.

Table 4: Assessment and Outcomes Matrix

	Competency	Movement	Experience	Accomplishment	Integrated Perf.
Reading (Participation)	X				
Quiz (Participation)	X				
Active Lectures (Participation)	X				
Annotated Code	X				
Peer Assessment	X				
Short Response		X	X	X	
Discovery		X	X	X	
Final Project			X		X

9. COURSE GRADING SYSTEM

The course grade will be determined with the following weight for class assignments:

- *Participation Activities (10%)*
  - *These activities require quick assessment and feedback. You will receive credit for completing these activities.*
- *Annotated Code Activities (5%)*
  - *In these activities you be asked to only submit a script or code file that contains comments and demonstrates active exploration of the objectives within the activity.*
- *Peer Assessment Activities (5%)*
  - *Some activities will require you to assess the work of your fellow students. In these activities, your performance will be based on your work assessed by the instructor, your feedback to peers, and your peers' assessment of your work.*
- *Short Response Activities (20%)*
  - *Many activities are structured such that you respond to a set of questions. You will receive credit both for completing these activities and for the depth and detail of your responses. We will attempt electronic submittal and feedback for these activities.*
- *Discovery Activities (30%)*
  - *These activities require you to build on knowledge and skills introduced to you in previous activities. These activities will be open ended and you will receive credit for completing these activities and for the creativity of your exploration.*
- *Final Project (30%)*
  - *A final independent structured analysis will be selected by the student from a set of open-ended questions devised by the instructor. This project will serve as the final assessment that the student has made progress in developing knowledge and skills in this class. The project is due during the final exam period, where students will make a brief presentation on their results to the class.*

10.FINAL COURSE OUTLINE

Week 1			
Schedule	*Title*	*Time*	*AM*
In Class	Lecture 0: Introduction to Course	50	***
Out of Class	Activity 1: Principles of Graphics	50	SR

Title	*Time*	*AM*
Lecture 1: Principles of Graphics	55	***
Activity 3: An Experiment in Graphical Perception	55	P
Activity 4: Critiquing a Graphic for Graphicay	70	SR

	Activity 2: Overview of the Datasets Available for the Class	60	QZ	Activity 5: Setting up Your Accounts	20	QZ

Week 2						
Schedule	*Title*	*Time*	*AM*	*Schedule*	*Time*	*AM*
In Class	Activity 6: Creating a Simple Database – An Excel Straw man	110	P	Lecture 2: Database and SQL	25	***
				Activity 8: Creating a Simple Database – Now with PostgreSQL	85	P
Out of Class	Activity 7: Overview of SQL	60	AC	Activity 9: Simple SQL	25	SR
				Activity 10: An Introduction to R and R Studio	30	QZ

Week 3						
Schedule	*Title*	*Time*	*AM*	*Schedule*	*Time*	*AM*
In Class	Lecture 3: An Introduction to R	20	***	Lecture 4: R Plots	25	***
	Activity 11: Setting Up R	20	P	Activity 15: Learning Some Simple Plotting Features of R	55	AC
	Activity 12: A Starting Point – Some Simple R	70	AC	Activity 16: Your First Advanced Plot	30	AC
Out of Class	Activity 13: Reading in Data Files	15	AC	*FINISH: Activity 16: Your First Advanced Plot*	60	
	Activity 14: R Plots	50	QZ			

Week 4						
Schedule	*Title*	*Time*	*AM*	*Schedule*	*Time*	*AM*
In Class	Activity 17: Code Sharing	50	PA	Lecture 5: Analysis and Programming	20	***

	Peer Feedback on Activity 4.	60	PA	Activity 18: Thinking like a Computer – Pseudo-coding	30	P
				Activity 19: Write Your Own Function	60	D
Out of Class						

Week 5						
Schedule	*Title*	*Time*	*AM*	*Schedule*	*Time*	*AM*
In Class	Activity 20: Connecting to the Class Database via RODBC and PostgreSQL Drivers	20	P	Peer Sharing of Activity 18	50	PA
	Activity 21: Using R with PostgreSQL	60	P	Activity 23: Working with Time in R	60	D
	Lecture 6 R Packages	30	***			
Out of Class	Activity 22: Packages	60	PA	Activity 24: Interactive Review of Basic Statistics Using R	40	P
	Activity 5: Basic Charts for Single Discrete Variable	60	QZ	Activity 25: Exploratory Plots for the Distribution of a Single Discrete Variable	50	QZ

Week 6						
Schedule	*Title*	*Time*	*AM*	*Schedule*	*Time*	*AM*
In Class	Lecture 7 Types of Data	25	**	Lecture 9 Histograms and KDE	25	***
	Activity 26: Exploring Single Discrete Variable Plots	85	SR	Activity 28: Kernel Density Estimates and Histograms	85	SR
Out of Class	Activity 27: Probability Distributions	50	P			

Week 7						
Schedule	*Title*	*Time*	*AM*	*Schedule*	*Time*	*AM*

In Class	Lecture 9 Normal Distribution	20	***	Activity 32: Introduce Research Topics	45	SR
	Activity 29: Diagnosing a Distribution	90	SR	Activity 33: Advanced Multivariate Displays and Diagnostics	65	SR
Out of Class	Activity 30: Depicting the Distribution Involving Discrete Variables	50	QZ	Activity 34: One and Two Sample Tests	50	QZ
	Activity 31: Depicting the Distribution of Two Continuous Variables	50	QZ			

Week 8						
Schedule	*Title*	*Time*		*Schedule*	*Time*	
In Class	Lecture 10 Review of Some Analysis Methods	20	***	Lecture 12: Approaches to Data Mining	30	***
	Activity 35: Exploring Confidence Intervals and Simple Hypothesis Testing	90	SR	Tutorial: Data Mining and the Use of Plyr Package	80	D
	Activity 36: An Application of Hypothesis Testing	70	D			
Out of Class	Activity 37: Visualizing Multivariate Data	70	QZ			

Week 9						
Schedule	*Title*	*Time*		*Schedule*	*Time*	
In Class	Lecture 11: Random Sampling	20	***	Final Projects		
	Activity 39: Introduction to Simulation Analysis	90	D			
Out of Class						

Week 10	11/27/2012		
Schedule	*Title*	*Time*	
In Class	Final Projects		
Out of Class			

11/29/2012		
Schedule	*Time*	
Final Projects		