

The Hong Kong University of Science and Technology

UG Course Syllabus

Introduction to data analytics for smart transportation systems

CIVL 4610

3 Credits

Prerequisites: CIVL 1121 OR COMP 1023 OR COMP 1021 OR COMP 1022P OR COMP 2011 OR COMP 2012H

Corequisite: CIVL3610

Name: Sisi Jian

Email: cesjian@ust.hk

Office Hours: By email appointments

Course Description

This course covers the role of stochasticity in transport systems and the methods used to account for this within transport infrastructure assessment, with a particular focus on the application of data analysis methods. The course introduces how to analyze the performance of public transport systems and road network using classic queuing theory and travel time reliability concepts. The course will complement skills learnt in the other transport courses to provide a well-rounded knowledge of smart transport planning and management. The focus is on the application of transport models in real world settings using real data. Students have the opportunity to work with large open-source data in two experiential-learning projects. The course also develops skills for working with data and managing collaborative projects.

Intended Learning Outcomes (ILOs)

This course contributes to the following program learning outcomes:

ILO1. Obtain fundamental knowledge in mathematics and science

Students will learn the fundamental knowledge of queuing theory and reliability, as well as stochastic characteristics of transport systems.

ILO4. Apply modern engineering tools

Students will learn how to work with large open-source data in Python, and learn how to use visualization tools and data mining techniques.

ILO5. Formulate problems and propose feasible solutions

Students will learn how to design a research question, methodology and data approach for a real problem. Students will apply project design and data analysis methods to real problems with real data.

ILO8. Obtain in-depth knowledge in at least one specialized area

Students will learn in-depth knowledge in queuing theory and performance evaluation of urban transportation systems

ILO9. Communicate ideas effectively and able to work in teams

Students will work in teams in the second part of the course and work on a group project. Students will also present their work in a poster showcase in the last week.

ILO12. Stay abreast of contemporary issues

This course focuses on the application of transport models in real world settings. Students will work on real data obtained from either government website or APIs, and utilize transport models to evaluate existing public transport and road network performance, identify current issues, and propose transport policy implications.

Assessment and Grading

This course will be assessed using criterion-referencing and grades will not be assigned using a curve. Detailed rubrics for each assignment are provided below, outlining the criteria used for evaluation.

Assessments:

Assessment Task	Contribution to Overall Course grade (%)	Due date
Individual Assignment 1	20%	Week 3
Individual Assignment 2	20%	Week 5
Group Project Part 1	10%	Week 8
Group Project Part 2	15%	Week 10
Group Project Part 3	25%	Week 12
Group Project Part 4	10%	Week 13

Mapping of Course ILOs to Assessment Tasks

Assessed Task	Mapped ILOs	Explanation
Individual Assignment	ILO1, ILO4, ILO5, ILO8, ILO12	This task assesses students' ability to understand and apply travel time reliability and queuing theory concepts (ILO1), work with large open-source data in Python (ILO4), design a research question, methodology and data approach for a real problem (ILO5), evaluate the performance of urban transportation systems (ILO8), and utilize transport models to identify current issues, and propose transport policy implications (ILO12).
Group Project	ILO4, ILO5, ILO8, ILO9, ILO12	This task assesses students' ability to work with large open-source data in Python (ILO4), design a research question, methodology and data approach for a real problem (ILO5),

		evaluate the performance of public transportation systems (ILO8), work in a team (ILO9), and utilize transport models to identify current issues, and propose transport policy implications (ILO12).
--	--	--

Final Grade Descriptors:

Grades	Short Description	Elaboration on subject grading description
A	Excellent Performance	Demonstrates a comprehensive grasp of subject matter, expertise in problem-solving, and significant creativity in thinking. Exhibits a high capacity for scholarship and collaboration, going beyond core requirements to achieve learning goals.
B	Good Performance	Shows good knowledge and understanding of the main subject matter, competence in problem-solving, and the ability to analyze and evaluate issues. Displays high motivation to learn and the ability to work effectively with others.
C	Satisfactory Performance	Possesses adequate knowledge of core subject matter, competence in dealing with familiar problems, and some capacity for analysis and critical thinking. Shows persistence and effort to achieve broadly defined learning goals.
D	Marginal Pass	Has threshold knowledge of core subject matter, potential to achieve key professional skills, and the ability to make basic judgments. Benefits from the course and has the potential to develop in the discipline.
F	Fail	Demonstrates insufficient understanding of the subject matter and lacks the necessary problem-solving skills. Shows limited ability to think critically or analytically and exhibits minimal effort towards achieving learning goals. Does not meet the threshold requirements for professional practice or development in the discipline.

Student Rubrics:

Individual Assignment 1 (20 individual points)

4 allocated Category with 5 points for each. Since each student may choose different components, the grading for this assignment will be based on a deduction system.

Category	Component	Content	Deducted points
Category 1	Component 1	Lack of data description (like time range, number of observation)	1.6 ~ 2
	Component 2	Insufficient dataset description: Lacked detailed summary (e.g., geographic extent) and merely listed data fields	0.5
	Component 3	Uncompleted figure title	0.5
		Meaningless figure (number of observations): e.g. Pie chart lacks meaningful categorization (e.g., by road type) to represent proportions	1 ~ 1.6
Category 2	Component 4	Incorrect use of flowchart cell	0.5
		Incorrect pseudocode form	0.8 ~ 1

		key steps lack explanation and visualization (pseudocode/flowchart)	0.8 ~ 1
	Component 5	Uncompleted figure title	0.5
		Inappropriate figure type	0.5 ~ 1
Category 3	Component 6	Meaningless analysis	1 ~ 2
	Component 7	Meaningless mapping: e.g. the color-coding system for routes/endpoints lacks a clarifying legend	1 ~ 2
	Component 8	Inappropriate expression / Meaningless figure	1 ~ 2
Category 4	Component 9	Without segmentation (only calculate the reliability of each road)	0.6 ~ 1
		Without or with meaningless analysis	1 ~ 1.6
Lack of component			3

Individual Assignment 2 (20 individual points)

Component	Content	Mark
Component 1	1. Explain the motivation to employ the model	1
	2. Explain the independent variables selected, and the reason to choose them	1
	3. Propose priori expectations of the model results	1
	4. Discuss strengths and weaknesses of your approach	1
Component 2	1. Table presentation of the results - Sample size, number of observations used to fit/train the model a) Model performance (R2, chi-squared style metrics, likelihood-based measures, ...) b) Estimated values for the coefficients and corresponding significance of independent variables - (comparisons among different models)	2
	2. Text analysis of the results a) how well the model describes the data b) which coefficients / independent variables are statistically significant	1
	3. Clear, informative and concise presentation - Clear, understandable notations and captions - Tidy format - Quantitative indicators to support your description - Only present results that contribute meaningfully to your reader	1
Component 3	1. Reserve a validation set for model assessment	0.5
	2. Text description of how well the model captures the data - Discuss any concerns that emerge in the validation - (comparisons among different models)	2
	3. Visualization of the validation results	1
	4. Clear, informative and concise presentation	0.5

Component	Content	Mark
Component 4	1. A clear response to the research question - a clear connection between the technical information (in components 2 and 3) and the research question	2
	2. An indication of confidence in the result (conclusive or marginal)	1
	3. Clear, informative and concise presentation	1
Component 5	1. Identify limitations of the project - Shortcomings in the data set and methodology - Issues that emerged from the model results	2
	2. Describe the next step improvements corresponding to the limitations	2

Group project part 1 (10 group points)

Component	Content	Mark
Q1	1. The question must be clear and specific . 2. The question must be explicitly related to queuing theory . 3. The question must be relevant to public transport scenarios . 4. The research scope proposed in the question must be practical and feasible .	2
Q2	1. The methodology must be fully consistent with the research question in Q1. 2. Clear and detailed procedures for obtaining all variables (e.g., data sources, measurement methods, calculation approaches) must be provided. 3. Performance metrics for evaluating the research outcomes must be clearly defined and justified.	2
Q3	The content and analysis of Q3 must be consistent with the methodology designed in Q2 , without logical contradictions or deviations.	1
Q4	The content and analysis of Q4 must be consistent with the methodology designed in Q2 , without logical contradictions or deviations.	1
Q5	The content and analysis of Q5 must be consistent with the methodology designed in Q2 , without logical contradictions or deviations.	1
Q6	The option " no aggregation " must be selected and justified with valid reasons relevant to the research context.	1
Task Allocation	1. The task allocation must cover all project components specified in the instruction document.	1

Component	Content	Mark
	2. The allocation plan must strictly comply with all requirements and guidelines outlined in the instruction document.	
Gantt Chart	1. The Gantt chart must be completely consistent with the task allocation plan (including task sequence, duration, and responsible parties). 2. The presentation task must be explicitly included in the Gantt chart with clear timeline and milestones.	1

Group project part 2 (15 individual points)

Students choose two components out of nine, with 7.5 points allocated to each. Since each student may choose different components, the grading for this assignment will be based on a deduction system.

Category	Component	Content	Deducted points
Category 1	Component 1	Lack or mistake of data description (like time range, number of observations, as well as the meaning of fraction).	0.5 ~ 2
		Lack of the description of other general attributes related to your group’s research topic.	0.5 ~ 1
	Component 2	Insufficient dataset description: Lacked detailed summary (e.g., geographic extent) and merely listed data fields	0.5
Category 2	Component 3	Lack of the relevance and applicability of the sources to your research topic.	0.5 ~ 2
	Component 4		
Category 3	Component 5	Lack of a clear summary of key data limitations, such as unmatched tap records, unreasonable trip durations, outliers, or the influence of public holidays/special events.	0.5 ~ 2
	Component 6	Incorrect use of flowchart cell	0.5
		Incorrect pseudocode form	0.5 ~ 1
		key steps lack explanation and visualization (pseudocode/flowchart)	0.5 ~ 1
Category 4	Component 7	Uncompleted figure title	0.5
		Inappropriate figure type	0.5 ~ 1
	Component 8 (compulsory)	Uncompleted figure title	0.5 ~ 1
		Unable to be clearly distinguished	0.5 ~ 2
Lack of component			3

Group project part 3 (15 individual points + 10 group points)

a) Personal: 3 allocated components with 5 points for each

Component	Content	Mark
Component 1	1. Indicate what model(s) you use	1
	2. Explain the motivation to employ the model (why appropriate for your question)	2

Component	Content	Mark
	3. Propose priori expectations of the model results	1
	4. Discuss strengths and weaknesses of your approach	1
Component 2	1. Including key assumptions Examples: - On data accuracy - On data supplement (e.g. bus waiting time) - On tourist behavior and system regularity - On model - Other necessary assumptions	2
	2. Impact analysis of each assumption	3
Component 3	1. Table presentation of model attributes and results - Sample size, number of observations used to fit/train the model a) Model performance (R2, chi-squared style metrics, likelihood-based measures, ...) b) Estimated values for the coefficients and corresponding significance of independent variables - (comparisons among different models)	3
	2. Text analysis of the results a) how well the model describes the data b) which coefficients / independent variables are statistically significant	2
Component 4	1. Reserve a validation set for model assessment (no need if dataset is small)	1
	2. Text description of how well the model captures the data - Discuss any concerns that emerge in the validation - (comparisons among different models)	2
	3. Visualization of the validation results	2
Component 5	1. A clear response to the research question - a clear connection between the technical information (in components 3 and 4) and the research question	3
	2. An indication of confidence in the result (conclusive or marginal)	2
Component 6	1. Motivation of the ML model: why it is suitable for your research question	1
	2. ML model configuration: algorithm, hyperparameters, metric	2
	3. Validation results of the ML model	2
Component 7	1. Visualization of the model performance comparison	2

Component	Content	Mark
	2. Text description of model performance comparison - What the ML model reveals - Where the two models agree and disagree	2
	3. Direction for improving the queuing model in future work	1
Component 8	1. Identify stakeholders	1
	2. How does the issue you've addressed impact each stakeholder group?	2
	3. How your results and findings support these	2
Component 9	1. Identify limitations of the project - Shortcomings in the data set and methodology - Issues that emerged from the model results	3
	2. Describe the next step improvements corresponding to the limitations	2

b) Group: 10 points for methodology correctness and report consistency

- Methodology correctness (6 points)
- Report consistency (4 points): Do your selected components in the report answer the research question step-by-step? Is any component useless or lacked for a complete story?

Group project part 4 (10 group points)

Component	Content	Mark
Poster content and message	1. Clear	1
	2. Concise	1
	3. Completeness	2
Poster design and layout	1. Visual attractiveness	1
	2. Logical and easy to read	1
	3. Neatness	1
Presentation	The performance when students answer questions. These points are lost if the student is absent.	3

Course AI Policy

The use of Generative AI is permitted and requested to assist students with brainstorming, drafting, and writing their reports.

Communication and Feedback

Assessment marks for individual assessed tasks will be communicated via Canvas within two weeks of submission. Students who have further questions about the mark should consult the instructor within five working days after the mark is received.

Required Texts and Materials

- Roess, Roger P., Elene S. Prassas, William R. McShane. Traffic Engineering. Third Edition, Upper Saddle River: Pearson Prentice Hall, 2004 (ISBN 0-13-142471-8)
- Vukan Vuchic. Urban Transit Operations, Planning and Economics – John Wiley & Sons, 2005;
- Daganzo, C. Fundamentals of Transportation and Traffic Operations, Pergamon-Elsevier, Oxford, U.K. (1997)
- de Neufville, Richard. "Applied Systems Analysis - Engineering Planning and Technology Management", McGraw Hill, 1990.
- Hall, W. Randolph. "Queueing Methods - For Services and Manufacturing", Prentice Hall, 1991.
- Ravindran, A., Phillips, Don T. and Solberg, James J. "Operations Research - Principles and Practice", John Wiley and Sons, 1987.

Academic Integrity

Students are expected to adhere to the university's academic integrity policy. Students are expected to uphold HKUST's Academic Honor Code and to maintain the highest standards of academic integrity. The University has zero tolerance of academic misconduct. Please refer to [Academic Integrity | HKUST – Academic Registry](#) for the University's definition of plagiarism and ways to avoid cheating and plagiarism.