Birla Institute of Technology & Science, Pilani
Hyderabad Campus

ACADEMIC – GRADUATE STUDIES AND RESEARCH DIVISION
BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI-HYDERABAD CAMPUS
FIRST SEMESTER 2021-2022
Course Handout Part II

Date: 12/08/2021

In addition to part I (General Handout for all courses appended to the time table) this portion gives further specific details regarding the course.

Course No.: CE G572

Course Title: Transportation Data Analytics

Instructor-in-charge: Prasanta Sahu (prasanta.sahu@hyderabad.bits-pilani.ac.in)
Office: D-327

Description : Research Design Concepts in transportation, Data collection and analysis techniques. Probability Distributions, Sampling and Measurement, Interval Estimation, Hypothesis Testing, Analysis of Variance. Simple, Multiple and Time-series Regression, Dynamic Regression Models, Structural Equation Models, Count Data Models. Supervised Learning Methods: Tree-based Methods and Support Vector Machines. Clustering Methods: Dimensionality Reduction (Principal Component Analysis, Independent Component Analysis), Clustering Methods: Hierarchical clustering, K-Means Clustering, Mean Shift Clustering, Density-based Clustering. Data analysis and modelling using R-Studio. Case Studies –Traffic Operations, Pavement Materials and Quality Control, Transportation Planning, Public Transit and Road safety and Highway Geometric Design.

1. Scope and objectives of the course:

Scope: The course will present a number of model-estimation methods that are used in transportation data analysis and other subject areas that deal with data analysis. Transportation agencies spend millions of rupees conducting research to improve their ability to plan, design, construct, maintain, and operate the transportation system. These research projects need to use appropriate experimental designs or data analysis techniques, thereby decreasing the data collection costs and increasing the likelihood of project success. The course aims to develop fundamental skills needed for transportation projects by combining statistical methods with required computer coding skills to explore, describe,

model and test datasets. The objective is to provide students with a generic background in the application of various statistical and econometric analysis techniques and to provide new ideas for analyzing transportation data in research. Econometric methods are vital to travel demand model development, travel behavior analysis, traffic safety studies, and economic decision making. The students will be able to utilize various model development techniques discussed in the class in a number of applications such as (1) statistical model development, (2) regression model estimation with examining various properties, (3) cross-sectional and panel data analysis, (4) ordinary least squares and maximum likelihood estimation, (5) time series analysis, (6) cluster analysis, (7) simultaneous model estimation techniques, etc.

Course Outcome: At the end of this course, the students are expected to develop ability to:
1. Develop an understanding of transportation data collection methods.
2. Design transportation data collection experiments.
3. Demonstrate understanding of statistical modeling in transportation engineering applications
4. Develop Regression models to properly analyze the relationships between different variables in transportation systems
5. Apply the concepts learned in the class using real world data and will also learn econometric model development in softwares such as R, Matlab, Stata, SPSS, Biogeme, and Limdep.

Student Learning Outcomes (SLOs) assessed in this course – **(a), (b), (c), (e), (h), (i), (j),** and **(k).**

2. Textbook(s):

Text Book (TB)

- **T1:** Washington, S., M. Karlaftis, and F. Mannering (2011) Statistical and econometric methods for transportation data analysis, Second Edition, Chapman & Hall/CRC, Boca Raton, FL, ISBN:142008285X

- **T2:** Gujarati, D. N. and Sangeetha (2011). Basic Econometrics. 6th Edition. McGraw-Hill/Irwin, ISBN: 0073375772

- **T3:** Makridakis, S.G., Wheelwright, S. C., Hyndman, R. J. (2016). Forecasting: Methods and Applications. 3rd Edition. Wiley Publications, ISBN: 978-0-471-53233-0

Reference Books (RB)

- **R1:** Greene. W. (2010). Econometric Analysis. Pearson, ISBN0-13-139538-6

- **R2:** Wooldridge, J., M. (2012). Introductory Econometrics: A Modern Approach. 5th Edition, Cengage Publications, ISBN: 1-111-53104-8

- **R3:** Statistics for Engineering and the Sciences (6th Edition) by Mendenhall and Sincich, Taylor & Francis Group. CRC Press 2016

- **R4:** R in Action: Data Analysis and Graphics with R (2nd Edition) by Robert Kabacoff, Manning Publications
- **R5:** Multivariate Analysis – I: Published by STHDA (http://www.sthda.com), Alboukadel Kassambara

Lecture wise Course Plan

Lecture No.	Topics Covered	Learning objectives	Reference to TB, RB	SLO*
1	Developing Effective Experiment Design and Data Analysis Plans	To understand the concepts of experiment designs for transportation data analysis	*T1:Ch.1; R1: Ch.1 ; R2: Ch. 1*	**a, c**
2-3	Some Questions and Answers About Experiment Design	To understand a question-based approach to developing an appropriate experiment	*T1: Ch.1 ; T3: Ch.2; R2: Ch.2*	**a,c,e**
4-6	Examples of Effective Experiment Design	To learn different types of transportation experiments.	*T2: Ch8; T3: Ch.1*	**a,b,j**
7	Sample, population, variables	To categorize correctly random variables considering samples and populations	*T1: Ch.2; T2:Ch.1; T3: Ch.2*	**a,c,e**
8- 9	Describing data and Case Study Examples	To analyse correctly summary statistics and graphs to understand important data sample features	*T1: Ch.3 ; R1: Ch.3 ; R2: Ch.1*	**a,b,j,k**
10 - 12	Probability Distributions Survey instrument design; Sampling procedures and Sample size estimation	To explain clearly how to estimate the expected value of a discrete random variable	*T1: Ch.5 ; R1: Ch.4 ; R2: Ch.2 ; R3: Ch.5*	
13-23	Regression analysis; Linear regression technique and related statistical parameters; Development of regression models from field datasets; Category analysis; Panel Data Regression; Temporal and geographical stability	To be able to develop various analytical models; To calibrate multiple linear regression equations	*T1: Ch.3-5; T2:Ch.2-9; T3: Ch.4 and 5*	**a,b,e,j,k**
23-24	Case Study Examples and Modelling with R	To be familiar with econometric and statistical packages such as R Studio	*R4: Ch.4-6*	**a,b,e,k**
25-32	Time-series Regression, Dynamic Regression Models, Structural Equation Models, Count Data Models	To understand clearly the limitations of a SLR/MLR model for estimation and prediction; To learn how to overcome SLR or MLR models limitation with DR and SEM model	*T1: Ch.7, 10 and 11; T2: Ch.20; T3:Ch.3,4 and 7; R1: Ch. 8; R2: Ch.7*	**a,b,e,k**

33-35	Case Studies Examples and Modelling with R	To develop, analyse and interpret the model parameters using R Studio	*R4: Ch.4-6*	**a,b,c,e.k**
35-40	Clustering Methods: Hierarchical clustering, K-Means Clustering, Mean Shift Clustering, Density-based Clustering.	To understand the clustering methods and their usefulness in transportation problems	*R4: Ch. 11,13*	**c,e,i,j**
41-42	Case Studies Examples and Modelling with R	To solve transportation application problems using R Studio	*T1: Ch. 9, 10*	**a,f,h,i,j**

*Student Learning Outcomes (SLOs):

SLOs are outcomes (a) through (k) plus any additional outcomes that may be articulated by the program.

- (a) an ability to apply knowledge of mathematics, science and engineering
- (b) an ability to design and conduct experiments, as well as to analyze and interpret data
- (c) an ability to design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability
- (d) an ability to function on multidisciplinary teams
- (e) an ability to identify, formulate, and solve engineering problems
- (f) an understanding of professional and ethical responsibility
- (g) an ability to communicate effectively
- (h) the broad education necessary to understand the impact of engineering solutions in a global, economic, environmental, and societal context
- (i) a recognition of the need for, and an ability to engage in life-long learning
- (j) a knowledge of contemporary issues
- (k) an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice.

Evaluation Scheme

Ec. No.	Evaluation component	Duration	Weightage	Date, time	Nature of component
1.	Quiz (at least two)	45 Minutes	10%	To be announced in class	OB
2.	Assignments (at least 2)	-	15%	Continuous	OB
3.	Term Paper	-	15%	Continuous	OB
4.	Mid-semester exam	90 Minutes	25%	TBA	OB/CB
5.	Comprehensive Exam	2 Hours	35%	13/12 AN	OB/CB

Office Consultation Hour: To be announced in the class.

Notices:　　　　　All Notices concerning the course will be displayed on **CMS, Google Classroom and Notice Board** of Civil Engg. Department.

Make up policy:　　Makeup will be given only to the genuine cases with prior permission.

Academic Honesty and Integrity Policy: Academic honesty and integrity are to be maintained by all the students throughout the semester and no type of academic dishonesty is acceptable.

Instructor-in-charge

CE G572