UNIVERSITY OF THE FRASER VALLEY

ORIGINAL COURSE IMPLEMENTATION DATE: January 2012
REVISED COURSE IMPLEMENTATION DATE: September 2020
COURSE TO BE REVIEWED (six years after UEC approval)**:** February 2027
Course outline form version: 05/18/2018

OFFICIAL UNDERGRADUATE COURSE OUTLINE FORM

Note: The University reserves the right to amend course outlines as needed without notice.

Course Code and Number: STAT 271	**Number of Credits: 3** Course credit policy (105)
Course Full Title: Introduction to Data Analysis and Statistical Modeling **Course Short Title:** Data Analysis and Modeling	
Faculty: Faculty of Science	**Department (or program if no department):** Mathematics & Statistics

Calendar Description:

A practical course on the modelling and analysis of statistical data using statistical software. Topics include graphical presentation, linear and nonlinear regression, Poisson log-linear and logistic regression, design and analysis of experiments, survival time analysis, and time series analysis.

Note: Students with credit for MATH 271 cannot take this course for further credit.

Prerequisites (or NONE):	One of the following: STAT 104 with a B, STAT 106, or STAT 270.
Corequisites (if applicable, or NONE):	
Pre/corequisites (if applicable, or NONE):	

Antirequisite Courses *(Cannot be taken for additional credit.)*
Former course code/number: **MATH 271**
Cross-listed with:
Dual-listed with:
Equivalent course(s):
(If offered in the previous five years, antirequisite course(s) will be included in the calendar description as a note that students with credit for the antirequisite course(s) cannot take this course for further credit.)

Typical Structure of Instructional Hours

Lecture/seminar hours	
Tutorials/workshops	
Supervised laboratory hours	50
Experiential (field experience, practicum, internship, etc.)	
Supervised online activities	
Other contact hours:	
Total hours	**50**

Labs to be scheduled independent of lecture hours: ☒ No ☐ Yes

Special Topics *(Double-click on boxes to select.)*
This course is offered with different topics:
☒ No ☐ Yes *(If yes, topic will be recorded when offered.)*

Independent Study
If offered as an Independent Study course, this course may be repeated for further credit: *(If yes, topic will be recorded.)*
☒ No ☐ Yes, repeat(s) ☐ Yes, no limit

Transfer Credit
Transfer credit already exists: *(See **bctransferguide.ca**.)*
☐ No ☒ Yes
Submit outline for (re)articulation:
☒ No ☐ Yes *(If yes, fill in transfer credit form.)*

Grading System
☒ Letter Grades ☐ Credit/No Credit

Maximum enrolment (for information only): 36

Expected Frequency of Course Offerings: Every Year

Department / Program Head or Director: Ian Affleck	**Date approved:**	April 22, 2020
Faculty Council approval	**Date approved:**	September 11, 2020
Dean/Associate VP:	**Date approved:**	September 11, 2020
Campus-Wide Consultation (CWC)	**Date of posting:**	February 5, 2021
Undergraduate Education Committee (UEC) approval	**Date of meeting:**	February 26, 2021

Learning Outcomes:

Upon successful completion of this course, students will be able to:

1. Explore different types of statistical data graphically and interpret the properties of the data.
2. Propose appropriate family of statistical models for analyzing a data set.
3. Analyze the relationship through linear regression model, investigate the utility and the assumptions, propose a remedy if any assumption is violated, interpret the effectiveness of the fitted model, make estimation and prediction.
4. Analyze the relationship through a nonlinear regression model including logistic regression for binary data and log-linear Poisson model for count data.
5. Design and analyze single-factor and factorial experiments, give recommendation for obtaining optimum response variable, and identify the sample size.
6. Analyze survival time data through Cox Proportional hazards model, distinguish between censored and non-censored data, apply empirical approach of Kaplan-Maier to estimate the survival function, interpret the hazard and survival functions.
7. Interpret the components of a time series data, apply the regression approach to explain the trend and seasonal effect, use the moving average and weighted moving average for prediction.
8. Perform basic statistical computing using statistical software and interpret the outputs.

Prior Learning Assessment and Recognition (PLAR)

☒ Yes ☐ No, PLAR cannot be awarded for this course because

Typical Instructional Methods *(Guest lecturers, presentations, online instruction, field trips, etc.; may vary at department's discretion.)*

Lectures, hands-on-computer lab

NOTE: The following sections may vary by instructor. Please see course syllabus available from the instructor.

Typical Text(s) and Resource Materials *(If more space is required, download Supplemental Texts and Resource Materials form.)*

	Author (surname, initials)	Title (article, book, journal, etc.)	Current ed.	Publisher	Year
1.		Coursepack	☐		
2.			☐		
3.			☐		

Required Additional Supplies and Materials *(Software, hardware, tools, specialized clothing, etc.)*

Typical Evaluation Methods and Weighting

Final exam:	50%	Assignments:	10%	Field experience:	%	Portfolio:	%
Midterm exam:	30%	Project:	10%	Practicum:	%	Other:	%
Quizzes/tests:	%	Lab work:	%	Shop work:	%	Total:	100%

Details (if necessary):

The above percentages may vary among instructors and years. The final exam is comprehensive. Students must obtain at least 40% on the final exam in order to receive credit for this course.

Typical Course Content and Topics

1. Statistical data and graphical methods: Observational and experimental data, continuous data, count data, categorical data, contingency tables, life time data time series data, graphical methods for presentation of different type of data, searching the properties of data through graphs, searching for pattern and unusual observations.
2. Linear regression method: Simple and multiple linear regression models, fitting the linear regression models, diagnostics, Box-Cox transformation, significance test of regression, confidence interval and significance test for coefficients, correlation, practical interpretation of coefficients and R-squared, prediction, indicator variables, stepwise regression, Nonlinear regression method: Nonlinear regression models, logistic regression model, fitting logistic model, test of overall adequacy of the model, interpretation of the coefficients, confidence interval and significance test for coefficients, odds ratio and interpretation, log-linear Poisson model for count data, fitting log-linear model, inference concerning log-linear model fitting and coefficients, Bayesian approach for fitting logistic and log linear models.
3. Design and analysis of experiments: Single-factor and factorial experimental designs, ANOVA model, fitting ANOVA model, model adequacy checking, F-test for equality of factor level means, multiple comparisons, regression approach and contour plots, determining sample size, experiments with random effects and mixed effects.
4. Survival data analysis: Exponential and Weibull distributions, survival and hazard functions, censoring, empirical survival function, Kaplan-Maier estimator, fitting exponential and Weibull models.
5. Time-series data analysis: Components of a time series data, moving average method, weighted moving average method, linear trend, some specific nonlinear trend, seasonal index, autocorrelation and Durbin-Watson test.