

CPA STRATEGIC: DATA ANALYTICS FOR FINANCE

Aim

The aim of this module is to provide a comprehensive understanding of the theory and practice of data analysis in the Accounting and wider business domains. Students who complete this course should develop the skills to apply and interpret data-based initiatives that address real-world problems across many financial activities such financial accounting, management accounting, taxation, auditing and corporate finance. The course will leverage the wide use of Excel among accounting practitioners and develop students understanding of the Excel Data Analysis functions.

Data Analytics for Finance as an Integral Part of the Syllabus

The study of Management Accounting at Foundation Level and Performance Management at Professional are prerequisites for the study of Data Analytics for Finance. Statistical techniques introduced in earlier papers are applied to the methods of Data Analytics.

Syllabus and Subject Dependencies

LEVEL	SUBJECT					
CPA Foundation	Financial Accounting	Management Fundamentals	Taxation		Management Accounting	
CPA Professional	Financial Reporting	Corporate Law	Advanced Taxation	Audit & Assurance	Performance Management	Managerial Finance
CPA Strategic	Advanced Financial Reporting	Strategy & Leadership	Advanced Tax Strategy	Advanced Audit & Assurance	Data Analytics for Finance	Strategic Corporate Finance

Minimum Intended Learning Outcomes

On successful completion of this module, students will be able to:

- Discuss the strategic processes, benefits and challenges of applying Data Analysis techniques to improve decision making for competitive advantage
- Apply statistical techniques, including regression and time series, to business datasets and interpret statistical results for various user groups.
- Describe the key concepts underpinning various supervised and unsupervised machine learning techniques
- Interpret outputs from a variety of Data Analytics software tools, such as Excel, R and Python, and discuss the risks in making business recommendations from these outputs
- Evaluate the benefits and risks of utilising machine learning techniques in a variety of business financial activities.
- Assess data protection, data privacy and other ethical issues associated with the utilisation of large corporate datasets

Indicative Syllabus

		Competency Level
Strategic Data Management in Finance	- Challenges in Human Decision Making - Data Analytics Processes - Introduction to CRISP/DM - Financial Big Data for Competitive Advantage - Strategic Data Management in Finance - Management Challenges in Data Driven Environments - Developing a Data Driven Culture	3 3 3 3 3 2
Exploratory Data Analysis	- Data Preparation – Normalization, Binning, Sampling - Descriptive Statistics - Data Quality Issues - Missing Values, Outliers - Visualizing Relationships Between Features - Measuring Covariance and Correlation - Data Distributions and Confidence Intervals	3 3 3 3 3 3
Statistics for Big data	- Effect size, Statistical power and sample size - Effect of Variation - Hypothesis testing - Interpret outputs from statistical software to analyse patterns in accounting data for signalling unexpected fluctuations e.g. Fraud Analysis, detecting anomaly transactions etc.	3 3 3 3
Dimension reduction	- Factor Analysis - Principal Component Analysis (PCA) - Interpret outputs from statistical software to analyse reasons behind fluctuations in accounting data e.g. defining cost drivers in Activity Based Costing etc.	3 3 3
Regression and Big Data	- Simple Linear Regression - Correlation Coefficient - Calculation of Regression parameters - Overfitting - The Variance/Bias Trade-off - Collinearity - Regression Model accuracy - Residual Analysis - Interpret outputs from statistical software to interpret relationships in accounting data e.g. Price/Revenue relationship, comparing performance of different products, regions etc.	3 3 3 3 3 3 3 3
Forecasting Techniques and Application	- Time Series basics - Decomposition of Time Series - Seasonality, Linear Trend models - Smoothing models - Interpret outputs from statistical software to support prediction of accounting data e.g. Forecasting Revenue/Margin, Predicting Going Concern status, What-if analysis etc.	3 3 3 3 3

Indicative Syllabus contd.

		Competency Level
Machine Learning	- Supervised Learning - Classification - Tree-Based Methods - Decision Trees - Logistic Regression - K-Nearest neighbours - Applications in Finance - Interpretation of software model output	3
	- Unsupervised Learning - Measuring Similarity - Clustering - Neural Networks - Support Vector Machines - Applications in Finance	3
	- Evaluate ML techniques for analysis of high-volume accounting data - Classifying fraud risk factors in accounting transactions - Profiling customer buying and payment performance based on known or unknown attributes - Clustering accounting data for internal control procedure	3
Model Evaluation	Misclassification Rate on a Hold-out Test Set	3
	ROC Curves	3
	Confusion Matrix – Accuracy, Precision, Recall	3
Data Ethics & Legal Considerations	Data Ethics in Finance	2
	Data Legislation GDPR	2
	Data and Statistical reporting	2

Learning Guide

Strategic Data Management in Finance

- Identify and discuss the challenges in Human Decision Making, particularly relating to large datasets.
- Discuss the benefits and challenges of utilising CRISP/DM to manage Finance Data Analysis projects
- Identify and discuss the strategic benefits to be derived from Financial Big Data
- Identify and discuss the management challenges in leveraging the benefits of Big Data for strategic competitive advantage

Exploratory Data Analysis

- Discuss the process of data cleaning and preparation – e.g. Normalization, Binning, Sampling
- Apply and Evaluate key descriptive statistics, including Covariance and Correlation, in a data set for large business datasets
- Discuss solutions to overcome data quality issues in Data Analysis projects - missing values, outliers etc.
- Apply and Evaluate methods for visualizing relationships between features

Statistics for Big data

- Explain relationships between sample size, effect size, statistical power
- Describe and Evaluate measures of variation for large datasets
- Describe hypothesis testing and evaluate outputs from hypothesis tests performed using software such as Excel, R etc.
- Interpret outputs from statistical software to analyse patterns in accounting data for signaling unexpected fluctuations - fraud analysis, detection of anomaly transactions etc.

Learning Guide contd.

Dimension reduction

- Explain PCA and factor analysis and discuss its uses in the analysis of large financial datasets
 - Interpret outputs from statistical software to analyse reasons behind fluctuations in accounting data e.g. defining cost drivers in Activity Based Costing etc.
-

Regression and Big Data

- Apply simple linear regression to find the line of best fit
 - Calculate and explain the relevance of the Correlation Coefficient
 - Calculate regression parameters, describing the significance of each
 - Explain what is meant by overfitting and how the Variance/Bias Trade-off can be used to overcome this issue
 - Discuss the role of regression analysis in Financial Data Analysis and Forecasting
 - Interpret software outputs, including ANOVA and ANCOVA, to evaluate regression models
 - Interpret model accuracy from analysis of model residuals using tabular and graphical residuals plots
 - Interpret outputs from statistical software to interpret relationships in accounting data e.g. Price/Revenue relationship, comparing performance of different products, regions etc
-

Forecasting Techniques and Application

- Apply techniques to decompose additive and multiplicative time series and extract components such as trend, seasonality etc.
 - Discuss the role of time series analysis in Financial Data Analysis and Forecasting
 - Interpret outputs from statistical software to support prediction of accounting data e.g. Forecasting Revenue/ Margin, Predicting Going Concern status, What-if analysis etc.
-

Machine Learning

- Explain the fundamental techniques that underpin a range of Supervised Learning techniques such as Classification, Tree-Based Methods, Decision Trees, Logistic Regression, K-Nearest neighbours,
 - Explain the fundamental techniques that underpin a range of Unsupervised Learning techniques such as Clustering, Neural Networks, Support Vector Machines
 - Discuss the benefits and challenges of applying machine learning techniques in the financial domain e.g. for classifying fraud risk factors in accounting transactions, profiling customer buying and payment performance based on known or unknown attributes, clustering accounting data for internal control procedure
-

Model Evaluation

- Understand the application of ROC Curves in the evaluate of Machine Learning models
 - Explain concepts such as model Accuracy, Precision, Recall
 - Interpret Confusion Matrix and ROC curve outputs to evaluate the performance of a model
-

Data Ethics & Legal Considerations

- Identify and discuss the ethical issues surrounding the use of data analytics in finance
 - Demonstrate an understanding of Data Legislation GDPR and its impact on data analytics
 - Discuss the societal impacts of the increasing use of Data Analysis techniques in Finance and Business
-

Recommended Reading

The books identified below indicate the focus of this course. These books are indicative only. They are currently available free of charge from Springer's website

Excel Data Analysis - Modelling and Simulation

Author Hector Guerrero

Currently available to download free at

<https://link.springer.com/book/10.1007%2F978-3-030-01279-3>

Statistical Analysis and Data Display - An Intermediate Course with Examples in R

Authors - Richard M. Heiberger, Burt Holland

Currently available to download free at

<https://link.springer.com/book/10.1007%2F978-1-4939-2122-5>

The Data Science Design Manual

Authors - Steven S. Skiena

Currently available to download free at

<https://link.springer.com/book/10.1007%2F978-3-319-55444-0>