Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
MIT-WPU

COURSE STRUCTURE

Course Code	**AID0PM14A**				
Course Category	**Program Elective**				
Course Title	**Fundamentals of Data Analytics**				
Teaching Scheme	**Lectures**	**Tutorials**	**Laboratory / Practical**	**Project**	**Total**
Weekly load hours	**3**	**0**	**2**	**0**	**5**
Credits	**3**	**0**	**1**	**0**	**4**
Assessment Schema Code	**TL3**				

Prerequisites: Foundation of Data warehousing and Data Mining

Course Objectives: Experiential data analysis versus hypothesis testing
Knowing the set of data, Investigating classification variables ,investigating numerical variables investigating correlations between many variables ,Choosing compelling data subsets for additional analysis.

Course Outcomes:
After completion of this course students will be able to:

1. Understand what data is, how it is collected, the role of metadata in understanding a given set of data.
2. Work with different types of data.
3. Understand the statistical views in data analytics.
4. Understand the importance of data visualization to drive more effective business decisions.
5. Understand charts, graphs, and tools used for analytics and use them to gain valuable insights.

Course Contents:

Unit 1: Data Analytics Overview
Introduction, Data Analytics-Importance, Types of Data Analytics -Descriptive Analytics, Diagnostic Analytics, Predictive Analytics, Prescriptive Analytics, Data Analytics Benefits: Decision-making, Cost Reduction and other benefits, Applications of Data Analytics, Examples, Difference between Data Analysis & Data Analytics, Difference between Data Analyst and Data Scientist. Difference between Business Analysts, and Business Intelligence Analysts, role, responsibilities, and skillsets required to be a Data Analyst

Dean
Faculty of Engineering and Technology

Unit 2: Dealing with Different Types of Data

Introduction, Terminologies in Data Analytics, Text analytics, web analytics, marketing analytics, Types of Data –Data, numerical data and categorical data, structured and unstructured data, ordinal data and nominal data, data types, and data format, types of data repositories such as Databases, Data Warehouses, Data Marts, Data Lakes, and Data Pipelines

Unit 3: Basics of Statistics

Introduction, Types of statistics, Descriptive statistics, Inferential statistics, basic concepts in statistics - Population, Sample, Parameter, Statistic, and variable, different types of variables – dependent and independent, extraneous variable, continuous and discrete, Qualitative and Quantitative, concept of noise, Measures of Centre – mean, mode, median, Measures of variation – variance, standard deviation, range, Importance of statistics.

Unit 4: Data pre-processing

Importance of data quality, fundamentals of data privacy, what is dirty data? Data-cleaning tools and techniques, Common data-cleaning pitfalls, Data-cleaning features in spreadsheets, Optimize the data-cleaning process, documenting the cleaning process, Using SQL to clean data, Spreadsheets versus SQL, widely used SQL queries, Data Transformation - various encoding techniques for categorical data, Data Reduction - Attribute Subset Selection, Dimensionality Reduction

Unit 5: Data Visualization for Decision making

Introduction to Data Visualization, Understanding Data Visualization, Commonly Used Visualizations, Frequency Distribution Plot, Swarm Plot, Importance of Data Visualization, Data Visualization Tools, creating advanced charts and visualizations

Laboratory Exercises / Practical:

1. Using Excel - filter and sort data, LEN, BASIC Arithmetic
2. Using excel - LEFT, RIGHT, MID, IF, IFERROR, CHOOSE
3. Using Excel - FIND, ISNUMBER, ISTEXT, SUBSTITUTE,
4. Using Excel - TRIM, COUNTA, COUNTBLANK, COUNTIF
5. Using Excel – Date and Time functions
6. Using Excel – INDEX, MATCH
7. Remove duplicate, inaccurate data, and empty rows in your data.
8. VLOOKUP, HLOOKUP reference functions
9. Basic Statistics – min, mode, median, Standard Deviation, Variance,Pivot table
10. Encoding Categorical Variables
11. Plotting Datasets using various chart types
12. Case study of Healthcare Data Analytics

Learning Resources:

Dean
Faculty of Engineering and Technology

Text Books/ Reference Books:

1. Dr. Anil Maheshwari, Data Analytics Made Accessible, Publisher: Amazon.com Services LLC.
2. Trevor Hastie, Robert .Tibshirani, and Jerome Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction" Published by Springer.
3. Robert Keane , Data Analytics: Practical Guide to Leveraging the Power of Algorithms, Data Science, Data Mining, Statistics, Big Data, and Predictive Analysis to Improve Business, Work, and Life" Published by Independently published.

Supplemen Practical Tableautary Reading:

1. Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers and Keying Ye, Probability & Statistics for Engineers & Scientists (9th Edn.) Prentice Hall Inc.
2. John M. Chambers, Software for Data Analysis: Programming with R (Statistics and Computing), Springer.
3. Seema Acharya, Subhashini Chellappan, "Big Data and Analytics", Wiley Publication, first edition. Reprint in 2016

Web Resources:

Weblinks:

1. https://careerfoundry.com/en/tutorials/data-analytics-for-beginners/introduction-to-data-analytics/
2. https://intellipaat.com/blog/tutorial/data-analytics-tutorial/
3. https://www.tutorialspoint.com/data-analytics-business-intelligence/index.asp
4. https://www.youtube.com/watch?v=dPUiCk-FjaE

MOOCs: Online courses for self learning

Courses by NPTEL and MIT Open Courseware etc

1. http://nptel.ac.in/courses
2. https://swayam.gov.in/
3. https://www.coursera.org/professional-certificates/ibm-data-analyst
4. https://www.mooc-list.com/

Pedagogy:

1.Power point presentations
2.Videos
3.Demonstrations
Systematic use of group work and project-based learning.

Dean
Faculty of Engineering and Technology