Data Science -VI Sem

RGPV (Diploma Wing) Bhopal	SCHEME FOR LEARNING OUTCOME	Branch Code			Course Code		CO Code	LO Code	Format No. **4**
		C	0	4			1	1	

COURSE NAME	Data Science
CO-1 Description	**Explain the basic Concepts of Data Science.**
LO-1 Description	**Explain the basic concepts of Data Science and its applications.**

SCHEME OF STUDY

S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks
1.1	• Introduction to Data Science, • Evolution of Data Science • Data Science Roles Stages in a Data Science Project • Applications of Data Science in various fields	Traditional Lecture method + Handout	Teacher will explain the contents and provide handouts to students.	08		Handout	

SCHEME OF ASSESSMENT

S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal
	Paper pen test **(End Semester Exam)**	Students will be asked to explain data science basics,roles,data science project life cycle and applications of Data Science .	**10**	Test Paper	**External**

	ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)

RGPV (Diploma Wing) Bhopal	SCHEME FOR LEARNING OUTCOME	Branch Code			Course Code		CO Code	LO Code	Format No. **4**
		C	*0*	*4*			*1*	*2*	

COURSE NAME	Data Science
CO-1 Description	**Explain the basic Concepts of Data Science.**
LO -2Description	**Demonstrate the concepts of Data Mining.**

SCHEME OF STUDY							
S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks
1.2	Definitions of data mining Techniques of data mining Applications of data mining Data mining steps	Traditional Lecture method + Handout	Teacher will explain the contents and provide handouts to students.	08	-	Handout	

SCHEME OF ASSESSMENT				
S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal

Paper pen test (End Semester Exam)	A Student will be asked to define data mining, techniques and applications of data mining.	10	Test Paper	External

ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)

RGPV (Diploma Wing) Bhopal	SCHEME FOR LEARNING OUTCOME	Branch Code			Course Code		CO Code	LO Code	Format No. **4**
		C	0	4			1	3	

COURSE NAME	Data Science
CO-1 Description	**Explain the basic Concepts of Data Science.**
LO-3 Description	**Discuss Data Warehousing and online analytical processing**

SCHEME OF STUDY

S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks
1.3	• Definition • Architecture • Data warehouse models : enterprise warehouse, data mart, virtual warehouse	Traditional Lecture method	Teacher will explain the contents and provide a Lab manual to the students. Teacher will provide guided practice to apply pointers and structures in a given situation.	-	8	Handout	

	● Data cube : multidimensional data model, schemas (stars, snowflakes, fact constellations) ● OLAP operations						

<div align="center">

SCHEME OF ASSESSMENT

</div>

S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal
	Internal Paper pen test (Progressive test -I)	Students will be asked to explain data warehouse architecture & models, OLTP operations	**10**	Paper pen test	**Internal**

<div align="center">

ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)

</div>

RGPV (Diploma Wing) Bhopal	**SCHEME FOR LEARNING OUTCOME**	Branch Code			Course Code		CO Code	LO Code	Format No. **4**
		C	0	4			2	4	

COURSE NAME	**Data Science**
CO-2 Description	**Describe various statistical concepts.**
LO-4 Description	**Explain various Data objects and attribute types**

<div align="center">

SCHEME OF STUDY

</div>

S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks

| 2.1 | Data objects and attribute types: what is attributes, nominal, binary, ordinal, numerical, discrete and continuous attributes | Traditional Lecture method + Handout | Teacher will explain the contents and provide handout to students. | 06 | --- | Handout | |

SCHEME OF ASSESSMENT

S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal
	End Semester Exam	Student will be asked to identify different data types	10	Test Paper	External

ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)

RGPV (Diploma Wing) Bhopal	SCHEME FOR LEARNING OUTCOME	Branch Code			Course Code		CO Code	LO Code	Format No. **4**
		C	*0*	*4*			*2*	*5*	

COURSE NAME	Data Science

CO-2 Description	**Describe various statistical concepts.**

LO-5 Description	**Illustrate various Statistical techniques**

SCHEME OF STUDY							
S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks
2.2	Basic Statistical Descriptions of Data- measuring the central tendency (mean, median ,mode and percentile) , measuring the dispersion of data (range, quartiles, variance, standard	Traditional Lecture method + Handout	Teacher will explain the contents and provide handout to students.	-	08	Handout	

| | deviation, and interquartile range) | | | | | |

SCHEME OF ASSESSMENT

S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal
	End Term Exam	Students will be asked to calculate central tendency measures and dispersion of data for a given problem.	**10**	Test Paper	**External**

ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)

		SCHEME FOR LEARNING OUTCOME	Branch Code		Course Code		CO Code	LO Code	
RGPV (Diploma Wing) Bhopal			C	0	4		2	6	Format No. **4**
COURSE NAME	Data Science								

CO-2 Description	**Describe various statistical concepts.**
LO-6 Description	**Recall the concepts of probability**

SCHEME OF STUDY

S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks

2.3	Concept of probability Methods of assigning probability (classical, relative frequency and subjective) Probability terminology (experiment, event, elementary event, sample space, mutually exclusive events, independent events), collectively exhaustive events, complementary events)	Traditional Lecture method	Teacher will explain the contents and provide handout to students.	8		Handout	

SCHEME OF ASSESSMENT

S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal
	End Term Exam	Student will be asked to calculate probability for a given problem	**10**	**Test Paper**	**External**

RGPV (Diploma Wing) Bhopal	SCHEME FOR LEARNING OUTCOME	Branch Code		Course Code		CO Code	LO Code	Format No. 4
	C 0 4					2	7	

COURSE NAME	Data Science
CO-2 Description	Describe various statistical concepts.
LO-7 Description	Use Appropriate Python Libraries to Manage a Dataset.

SCHEME OF STUDY

S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks
2.4	Setup Python IDE (Jupyter, Pycharm, Google Colab (Online Platform), etc.). Importing Libraries (pandas, numpy, matplotlib) Reading Files (.csv, .xlsx, .pdf, Image, etc.). Managing the dataset (preprocessing dataset) Splitting dataset (Kaggle, UCI ML repository, OpenML, MNIST, etc.) into training and test test. Perform Data visualization.	Traditional Lecture method + Handout	Teacher will provide guided practice session to the students		8	Lab Manual	

SCHEME OF ASSESSMENT

S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal

Lab Assessment	Student will be asked to work on python to read, manage, split and visualize different data sets	10	Lab Manual/Computer	External

ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)

RGPV (Diploma Wing) Bhopal	SCHEME FOR LEARNING OUTCOME	Branch Code		Course Code		CO Code	LO Code	Format No. 4
		C	0	4		3	8	

COURSE NAME	Data Science
CO-3 Description	**Discuss the various Data Mining Techniques**
LO-8 Description	**Explain Association rule mining.**

SCHEME OF STUDY

S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks
3.1	Market basket analysis and its case study Frequent itemsets closed itemsets association rules apriori algorithm for association rules with example	Traditional Lecture method + Handout	Teacher will explain the contents and provide handout to students.	7	-	Hand out	

SCHEME OF ASSESSMENT

S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal

End semester Exam	Students will be asked to explain market basket analysis , itemsets and apriori algorithm. Students will be asked to solve Numerical problem based on apiori algorithm .	10	Test paper	External

ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)

RGPV (Diploma Wing) Bhopal	SCHEME FOR LEARNING OUTCOME	Branch Code			Course Code		CO Code	LO Code	Format No. **4**
		C	0	4			3	9	

COURSE NAME	Data Science
CO-3 Description	**Discuss the various Data Mining Techniques**
LO-9 Description	**Perform association rule mining on selected data set**

SCHEME OF STUDY

S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks
3.2	Select a data set from a repository Apply apriori algorithm to generate association rules Examine the validity of generated rules	Traditional Lecture method + Handout	Teacher will provide guided practice session to students	06	-	Lab Manual	

SCHEME OF ASSESSMENT

S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal

Lab Assessment	Student will be asked to read data and apply apriori algorithm to generate valid association rules.	10	Lab Manual/computer	Internal
	ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)			

RGPV (Diploma Wing) Bhopal	**SCHEME FOR LEARNING OUTCOME**	Branch Code			Course Code		CO Code	LO Code	Format No. **4**
		C	0	4			3	10	

COURSE NAME	**Data Science**
CO-3 Description	**Discuss the various Data Mining Techniques**
LO-10 Description	**Discuss Clustering techniques**

SCHEME OF STUDY

S. No.	**Learning Content**	**Teaching –Learning Method**	**Description of T-L Process**	**Teach Hrs.**	**Pract. /Tut Hrs.**	**LRs Required**	**Remarks**
3.3	Basic Concepts of Cluster Analysis and its case study Major Clustering Approaches-Partitioning Methods, Hierarchical Methods, Density-Based Methods k-mean clustering with example	Traditional Lecture method	Teacher will explain the contents and provide Lab Manual to students.	7		Handout	

SCHEME OF ASSESSMENT

S. No.	**Method of Assessment**	**Description of Assessment**	**Maximum Marks**	**Resources Required**	**External / Internal**

Internal Assessment (Term work)	Students will be asked to explain clustering and its types. Students will be able to solve numerical problems using k-mean clustering.	10	Test Paper/Quiz/presentation	Internal

ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)

RGPV (Diploma Wing) Bhopal	SCHEME FOR LEARNING OUTCOME	Branch Code			Course Code		CO Code	LO Code	Format No. **4**
		C	0	4			3	11	

COURSE NAME	Data Science
CO-3 Description	**Discuss the various Data Mining Techniques**
LO-11 Description	**Describe various Classification techniques.**

SCHEME OF STUDY

S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks
3.4	Case study of classification Basic classification approaches – decision tree induction, Bayes classification, k-Nearest neighbor classification	Traditional Lecture method + Handout	Teacher will explain the contents and provide handout to students.	7	-	Handout	

SCHEME OF ASSESSMENT

S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal

	Progressive Test -II	student will be asked to describe classification and their types	**10**	**Test paper**	**Internal**
	ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)				

RGPV (Diploma Wing) Bhopal		**SCHEME FOR LEARNING OUTCOME**	**Branch Code**		**Course Code**		**CO Code**	**LO Code**	Format No. **4**
			C	*0*	*4*		*3*	*12*	
COURSE NAME	Data Science								
CO-3 Description	**Discuss the various Data Mining Techniques**								
LO-12 Description	**Perform classification on selected data set**								
SCHEME OF STUDY									

S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks
3.5	Select a data from a repository Apply any classification on it Evaluation of classes	Traditional Lecture method + Handout	Teacher will provide guided practice session to students	-	6	Lab Manual	
SCHEME OF ASSESSMENT							

S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal
	Lab Assessment	Students will be asked to read the dataset and apply classification on it and also be able to evaluate the classes .	10	Lab Manual and computer	Internal

ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)

RGPV (Diploma Wing) Bhopal	SCHEME FOR LEARNING OUTCOME	Branch Code			Course Code		CO Code	LO Code	Format No. 4
		C	0	4			4	13	

COURSE NAME	Data Science
CO-4 Description	**Differentiate between correlation and regression .**
LO-13 Description	**Compare Correlation and regression technique.**

SCHEME OF STUDY

S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks
4.1	Definition of correlation Scatter diagram Definition of Regression. Linear Regression. Logistic Regression.	Traditional Lecture method + Handout	Teacher will explain the contents and provide handout to students.	8		Handout	

SCHEME OF ASSESSMENT

S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal
	End Term Exam	Students will be asked to define correlation & regression with its types.	10	**Test Paper**	**External**
		ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)			

RGPV (Diploma Wing) Bhopal	**SCHEME FOR LEARNING OUTCOME**	Branch Code			Course Code		CO Code	LO Code	Format No. **4**
		C	0	4			4	14	

COURSE NAME	Data Science
CO-4 Description	**Differentiate between correlation and regression.**
LO-14 Description	**Demonstrate correlation between variables of selected data set**

SCHEME OF STUDY

S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks
4.2	Find correlation between variables and analyse the resultant matrix Create a scatter graph between variables.	Traditional Lecture method + Handout	Teacher will provide guided practice session to students	-	5	Lab Manual	
			SCHEME OF ASSESSMENT				

S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal
	External Lab Assessment	Students will be asked to read data and generate correlation matrices. Students will be able to generate scatter graph and explain it.	10	Lab Manual and computer	External
		ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)			

RGPV (Diploma Wing) Bhopal	SCHEME FOR LEARNING OUTCOME	Branch Code			Course Code		CO Code	LO Code	Format No. **4**
		C	0	4			4	15	

COURSE NAME	Data Science
CO-4 Description	**Differentiate between correlation and regression.**
LO-15 Description	**Demonstrate regression between variables of selected data set**

<div align="center">SCHEME OF STUDY</div>

S. No.	Learning Content	Teaching –Learning Method	Description of T-L Process	Teach Hrs.	Pract. /Tut Hrs.	LRs Required	Remarks
4.3	Create linear regression model to predict value of a variable, Analyse the result	Traditional Lecture method + Handout	Teacher will provide guided practice session to students	-	5	Lab Manual	
		SCHEME OF ASSESSMENT					

S. No.	Method of Assessment	Description of Assessment	Maximum Marks	Resources Required	External / Internal
	Lab Assessment	Students will be asked to read the data and create a regression model for it.	10	**Lab Manual and computer**	**External**
ADDITIONAL INSTRUCTIONS FOR THE HOD/ FACULTY (IF ANY)					

RGPV (DIPLOMA WING) BHOPAL	OBE CURRICULUM FOR THE COURSE	FORMAT-3	Sheet No. 1/5
Branch	**CSE**	**Semester**	**VI**

Course Code		Course Name	Data Science		
Course Outcome - 1		**Explain the basic Concepts of Data Science.**		**Teach Hrs**	**Marks**
Learning Outcome 1		**Explain the basic concepts of Data Science and its applications.**		**08**	**10**
Contents		Introduction to Data Science, Evolution of Data Science Data Science Roles Stages in a Data Science Project Applications of Data Science in various fields			
Method of Assessment		External: End semester theory examination (Pen paper test).			
Learning Outcome 2		**Demonstrate the concepts of Data Mining.**		**08**	**10**
Contents		Definitions of data mining Techniques of data mining Applications of data mining Data mining steps			
Method of Assessment		External: End semester theory examination (Pen paper test).			
Learning Outcome 3		**Discuss Data Warehousing and online analytical processing**		**08**	**10**
Contents		Definition Architecture Data warehouse models : enterprise warehouse, data mart, virtual warehouse Data cube : multidimensional data model, schemas (stars, snowflakes, fact constellations) OLAP operations			
Method of		Internal: Mid semester theory examination (Pen paper test).			

Assessment	
RGPV (DIPLOMA WING) BHOPAL	**OBE CURRICULUM FOR THE COURSE** ‖ **FORMAT-3** ‖ **Sheet No. 2/5**

Branch	CSE			Semester	VI

Course Code		Course Name	Data Science		

Course Outcome - 2	**Describe various statistical concepts.**	Teach Hrs	Marks
Learning Outcome 4	**Explain various Data objects and attribute types**	6	10
Contents	Data objects and attribute types: what is attributes, nominal , binary , ordinal , numerical , discrete and continuous attributes		
Method of Assessment	External: End semester theory examination (Pen paper test).		
Learning Outcome 5	**Illustrate various Statistical techniques**	8	10
Contents	Basic Statistical Descriptions of Data- measuring the central tendency (mean, median ,mode and percentile) , measuring the dispersion of data (range, quartiles, variance, standard deviation, and interquartile range)		
Method of Assessment	External: End semester theory examination (Pen paper test).		
Learning Outcome 6	**Recall the concepts of probability**	8	10
Contents	Concept of probability Methods of assigning probability (classical, relative frequency and subjective) Probability terminology (experiment, event, elementary event, sample space, mutually exclusive events, independent events), collectively exhaustive events, complementary events)		
Method of Assessment	External: End semester theory examination (Pen paper test).		
Learning Outcome 7	**Use Appropriate Python Libraries to Manage a Dataset.**	08	10

Contents	Setup Python IDE (Jupyter, Pycharm, Google Colab (Online Platform), etc.). Importing Libraries (pandas, numpy, matplotlib) Reading Files (.csv, .xlsx, .pdf, Image, etc.). Managing the dataset (preprocessing dataset) Splitting dataset (Kaggle, UCI ML repository, OpenML, MNIST, etc.) into training and test test. Perform Data visualization.
Method of Assessment	External: Laboratory observation and viva voce.

RGPV (DIPLOMA WING) BHOPAL	OBE CURRICULUM FOR THE COURSE	FORMAT-3	Sheet No. 3/5

Branch	CSE		Semester	VI

Course Code		Course Name	Data Science

Course Outcome - 3	**Discuss the various Data Mining Techniques**	Teach Hrs	Marks
Learning Outcome 8	**Explain Association rule mining.**	07	10
Contents	Market basket analysis and its case study Frequent itemsets closed itemsets association rules apriori algorithm for association rules with example		
Method of Assessment	External: End semester theory examination (Pen paper test).		
Learning Outcome 9	**Perform association rule mining on selected data set**	06	10
Contents	Select a data set from a repository Apply apriori algorithm to generate association rules Examine the validity of generated rules		
Method of Assessment	Internal: Laboratory observation		
Learning Outcome 10	**Discuss Clustering techniques**	07	10

Contents	Basic Concepts of Cluster Analysis and its case study Major Clustering Approaches- Partitioning Methods, Hierarchical Methods, Density-Based Methods k-mean clustering with example		
Method of Assessment	Internal: Term work		
Learning Outcome 11	**Describe various Classification techniques.**	07	10
Contents	Case study of classification Basic classification approaches – decision tree induction, Bayes classification, k- Nearest neighbor classification		
Method of Assessment	Internal: Mid semester theory examination (Pen paper test).		
Learning Outcome 12	**Perform classification on selected data set**	06	10
Contents	Select a data from a repository Apply any classification on it Evaluation of classes		
Method of Assessment	Internal: Laboratory observation		

RGPV (DIPLOMA WING) BHOPAL	OBE CURRICULUM FOR THE COURSE	FORMAT-3	Sheet No. 4/5
Branch	**CSE**	**Semester**	**VI**

Course Code		**Course Name**	**Data Science**		
Course Outcome - 4		**Differentiate between correlation and regression**		**Teach Hrs**	**Marks**
Learning Outcome 13		**Compare Correlation and regression technique**		08	10
Contents		Definition of correlation Scatter diagram Definition of Regression. Linear Regression. Logistic Regression.			
Method of Assessment		External: End semester theory examination (Pen paper test).			

Learning Outcome 14	Demonstrate correlation between variables of selected data set	05	10
Contents	Find correlation between variables and analyse the resultant matrix Create a scatter graph between variables.		
Method of Assessment	External: Laboratory observation and viva voce.		
Learning Outcome 15	Demonstrate regression between variables of selected data set	05	10
Contents	Create linear regression model to predict value of a variable, Analyse the result		
Method of Assessment	External: Laboratory observation and viva voce.		

REFERENCE BOOKS:

S No	Title & Publication	Author
1	Data Mining: Concepts and Techniques, Elsevier	Jiawei Han, Micheline Kamber, Jian Pei
2	Applied statistics and probability for engineers	Douglas Montgomery
3	Mastering python for data science	Samir Madhavan
4	E-books/E-tools/Relevant software to be used as recommended by AICTE/NITTTR/RGPV	

Suggested practical list -
1. Import and summarize the various datasets using python.
2. Preprocess the dataset according to need (manage missing data, encode data etc.) using python.
3. Perform data visualization using python.
4. Apply association rule mining using inbuilt functions and analyse the result using python.
5. Apply classification using inbuilt functions and analyse the accuracy of classification using python.
6. Apply correlation and visualize using inbuilt function the result in scattered graphs using python.
7. Apply regression analysis using inbuilt function and interpret the result using python.