	Ahsanullah University of Science and Technology		
	Department of Computer Science and Engineering		
	SET A, Class Test #1, Fall 2023		
	Course Code: CSE 4261	Course Title: Data Analytics	
	Time: 15 Minutes	Date: 6/5/2024	Full Marks: 3+4+3
ID:		Name:	

1. What are the current challenges with Big Data?

2. A sample of 400 male students is found to have a mean height of 67.47 inches. Can it be reasonably regarded as a sample from a large population with a mean height of 67.39 inches and a standard deviation of 1.30 inches? Test at a 5% level of significance.

3. Write down the implementation steps of the ISO Map Algorithm.

	Ahsanullah University of Science and Technology		
	Department of Computer Science and Engineering		
	SET B, Class Test #1, Fall 2023		
	Course Code: CSE 4261	Course Title: Data Analytics	
	Time: 15 Minutes	Date: 6/5/2024	Full Marks: 3+4+3
ID:		Name:	

1. Write the difference between low-pass and high-pass filters.

2. From a random sample of 36 New York civil service personnel, the mean age and the sample standard deviation were found to be 40 years and 4.5 years respectively. Construct a 95 percent confidence interval for the mean age of civil servants in New York.

3. Write down the implementation steps of the Local Linear Embedding (LLE) Algorithm.

	Ahsanullah University of Science and Technology		
	Department of Computer Science and Engineering		
	SET C, Class Test #1, Fall 2023		
	Course Code: CSE 4261	Course Title: Data Analytics	
	Time: 15 Minutes	Date: 6/5/2024	Full Marks: 3+4+3
ID:		Name:	

1. How do you measure the Multicollinearity in a dataset?

2. The foreman of ABC Mining Company has estimated the average quantity of iron ore extracted to be 36.8 tons per shift and the sample standard deviation to be 2.8 tons per shift, based upon a random selection of 4 shifts. Construct a 90 percent confidence interval around this estimate.

3. Write down the implementation steps of the **Synthetic Minority Oversampling Technique (SMOTE)** Algorithm.

	Ahsanullah University of Science and Technology
	Department of Computer Science and Engineering
	SET D, Class Test #1, Fall 2023
	Course Code: CSE 4261	Course Title: Data Analytics
	Time: 15 Minutes	Date: 6/5/2024	Full Marks: 3+4+3
ID:	Name:

1. Write the difference between **bagging and boosting sampling techniques** in ensemble classifiers.

2. The following nine observations were drawn from a normal population:

   27 19 20 24 23 29 21 17 27

   Test the null hypothesis H0: $\mu = 26$ against the alternative hypothesis Ha: $\mu \neq 26$. At what level of significance can H0 be rejected?

3. Write down the working strategies of the **Butterworth filter** Algorithm.

	Ahsanullah University of Science and Technology
	Department of Computer Science and Engineering
	SET E, Class Test #1, Fall 2023
	Course Code: CSE 4261	Course Title: Data Analytics
	Time: 15 Minutes	Date: 6/5/2024	Full Marks: 3+4+3
ID:	Name:

1. Write the need for the analytical sandbox in Big data analytics.

2. A 10-year-old survey of CPAs (Certified Public Accountants) in the U.S. found that their average salary was $60,014. An accounting researcher would like to test whether this average has increased over the years. A sample of 125 CPAs produced a mean salary of $68,695. Also given that the population standard deviation $\sigma = \$10,530$. Level of significant=1%.

3. Write down the implementation steps of the **Local Linear Embedding (LLE)** Algorithm.

	Ahsanullah University of Science and Technology
	Department of Computer Science and Engineering
	SET F, Class Test #1, Fall 2023
	Course Code: CSE 4261	Course Title: Data Analytics
	Time: 15 Minutes	Date: 6/5/2024	Full Marks: 3+4+3
ID:	Name:

1. What are the six phases of the Data Analytics Life Cycle?

2. A researcher collects 25 examples of products and determines the percent salt in each. These 25 examples give a sample mean salt content of 40.24 and a sample standard deviation of 8.93. Compute a 90% confidence interval estimate of the true variance of the percentage water for this new process.

3. Write down the implementation steps of the **Local Maximum Fitting (LMF)** Algorithm.