Section	Description	Points	Grade Breakdown and Requirements			Weightage
		60	What 80-100% looks like	What 60-80% looks like	What <60% looks like	
Problem 1 - Define the problem and perform Exploratory Data Analysis	- Problem definition - Check shape, Data types, statistical summary - Univariate analysis - Bivariate analysis - Use appropriate visualizations to identify the patterns and insights - Key meaningful observations on individual variables and the relationship between variablesa	6	- Problem definition [0.5] - Check shape, Data types, statistical summary [0.5] - Use appropriate visualizations to identify the patterns and insights - Univariate Analysis [1.5] - Multivariate Analysis [2.5] - Key meaningful observations on individual variables and the relationship between variables [1]	- Problem definition - Check shape, Data types - Use appropriate visualizations to identify the patterns and insights - Univariate analysis done for some variables but not all - Multivariate analysis done for 2-3 combination of variables	- Problem definition - Some EDA done	10.00%
Problem 1 - Data Pre-processing	Prepare the data for modelling: - Outlier Detection(treat, if needed) - Feature Engineering / drop redundant features (if needed) - Encode the data - Train-test split	2	Prepare the data for modelling: - Outlier Detection(treat, if needed) [0.5] - Drop redundant variables [0.5] - Encode the data [0.5] - Train-test split [0.5]	- Missing value Treatment - Encode the data - Train-test split	- Encode the data - Train-test split	3.33%
Problem 1 - Model Building - Bagging	- Build a Bagging classifier - Build a Random forest classifier - Check the performance of the models across train and test set using different metrics and comment on the same	5	- Build a Bagging classifier [1.5] - Build a Random forest classifier [1.5] - Check the performance of the models across train and test set using different metrics and comment on the same [2]	- Build a Bagging classifier - Build a Random forest classifier - Check the performance of the models across train and test set using different metrics	- Build a Bagging classifier or Random forest classifier	8.33%
Problem 1 - Model Improvement - Bagging	- Try and improve the model performance by tuning the model (minimum 2 parameters to be tuned) - Bagging Classifier - Random Forest Classifier - Comment on model performance after tuning the model	7	- Try and improve the model performance by tuning the model (minimum 2 parameters to be tuned) - Bagging Classifier [3] - Random Forest Classifier [3] - Comment on model performance after tuning the model [1]	- Try and improve the model performance by tuning the model - Bagging Classifier - Random Forest Classifier	- Try and improve the model performance by tuning the model (done for one model)	11.67%
Problem 1 - Model Building - Boosting	- Build a Boosting classifier - Check the performance of the models across train and test set using different metrics and comment on the same Note: AdaBoost or GradientBoosting classifier can be built	3	- Build a Boosting classifier [1.5] - Check the performance of the models across train and test set using different metrics and comment on the same [1.5]	- Build a Boosting classifier - Check the performance of the models across train and test set using different metrics	- Build a Boosting classifier	5.00%
Problem 1 - Model Improvement - Boosting	- Try and improve the model performance by tuning the model (minimum 2 paratmeters to be tuned) - Comment on model performance after tuning the model	4	- Try and improve the model performance by tuning the model (minimum 2 parameters to be tuned) [3] - Comment on model performance after tuning the model [1]	- Try and improve the model performance by tuning the model (minimum 2 parameters to be tuned)	- Try and improve the model performance by tuning the model	6.67%
Problem 1 - Actionable Insights & Recommendations	- Compare all the models and choose the best model with proper rationale - Conclude with the key takeaways (actionable insights and recommendations) for the business	6	- Compare all the models and choose the best model with proper rationale [2] - Conclude with the key takeaways (actionable insights and recommendations) for the business - Actionable insights [1 x 2] - Recommendations [1 x 2]	- Comment on the importance of features based on the best model - Some actionable insights and recommendations mentioned	- Some actionable insights mentioned	10.00%

Problem 2 - Perform Exploratory Data Analysis and Text Pre-processing	- Perform exploratory data analysis
	  - Missing Value Checking and Treatment
	  - Feature Engineering
	  - Analysis of tweets
	  - Analysis of twitter activity
	  - Plot wordcloud
	
	- Text pre-processing
		9	- Perform exploratory data analysis
			  - Missing Value Checking and Treatment [0.5]
			  - Create columns for tweet year, month, day, and hour [0.5]
			  - Analysis of tweets (by likes, retweets in the presidential year) [1]
			  - Analysis of tweets (by characters, words, and sentiment) [1]
			  - Analysis of twitter activity (by hour and day) [1]
			
			- Text pre-processing
			  - Tokenization [1]
			  - Convert to lower case [0.5]
			  - Removal of punctuation [0.5]
			  - Removal of stopwords [1]
			  - Lemmatization [1]
			  - Plot wordcloud [1]
				- Perform exploratory data analysis
				  - Missing Value Checking and Treatment
				  - Create columns for tweet year, month, day, and hour
				  - Analysis of tweets (by likes, retweets in the presidential year)
				  - Analysis of tweets (by characters, words, and sentiment)
				
				- Text pre-processing
				  - Tokenization
				  - Convert to lower case
				  - Removal of punctuation
					- Perform exploratory data analysis
					  - Missing Value Checking and Treatment
					  - Some analysis of tweets
					
					- Some text pre-processing done
						15.00%

Problem 2 - Insight Generation	- Provide insights from the pre-processed data	3	- Summarise findings from Trump's tweet analysis about his usage of twitter [0.5 x 6]	- Summarise findings from Trump's tweet analysis about his usage of twitter (3-4 points mentioned)	- Summarise findings from Trump's tweet analysis about his usage of twitter (1-2 points mentioned)	5.00%
Business Report Quality	- Adhere to the business report checklist	6	Objective, guidance, and data description: 1 point Exclusion of code: 2 points Structure and readability: 1 point Rationale and logic: 1 point Visual clarity and referencing: 1 point	Objective, guidance, and data description Structure and readability Rationale and logic	Objective, guidance, and data description Rationale and logic	10.00%
Guided Project Deduction		9	Zero marks should be awarded in this section always as this is a deduction			15.00%