# NOTE:

**1. This Is just a template of the report from my part and you would rather need to customize this particular document to fit into context of your code, examination rules, lecturer presentation preference and proper examination rules.**

**2. Read and Run all your R codes and add all visualized diagrams and results to this document as demanded from the questions.**

**3. You might not have some packages I used that I used so you need to install it in Rstudio before loading it. i.e (install.package(“**packagename**”).**

**REPORT TEMPLATE:**

# ANALYSIS REPORT:

# PREDICTIVE MODELING OF H1N1 VACCINE UPTAKE

## **Introduction**

This analysis aims to explore and predict the likelihood of individuals receiving the seasonal H1N1 flu vaccine. The dataset used in this study contains various demographic and attitudinal factors related to vaccine uptake. We applied the Naive Bayes classification algorithm to predict vaccine uptake and evaluated the model's performance using various metrics.

## **Data Preprocessing**

The dataset was preprocessed to ensure compatibility with the Naive Bayes algorithm. Categorical variables were converted to factors, and the dataset was divided into training and test sets. Feature selection was performed using Cramer's V statistic to identify relevant predictor variables for the modeling process.

## **Naive Bayes Modeling**

We trained a Naive Bayes classification model using the training set. The goal was to predict whether an individual would receive the seasonal H1N1 flu vaccine based on various predictor variables. The model used a probabilistic approach to estimate the likelihood of vaccine uptake for each individual.

## **Model Evaluation**

### **Confusion Matrix:**

The confusion matrix provides insights into the performance of the Naive Bayes model. It shows the number of true positives, true negatives, false positives, and false negatives. From the confusion matrix, we calculated several performance metrics:

* **Overall Accuracy:** The model achieved an overall accuracy of X%, indicating the proportion of correct predictions among all predictions.
* **Specificity:** The specificity of the model was Y%, representing its ability to correctly identify non-vaccinated individuals.
* **Precision:** The precision of the model was Z%, indicating the proportion of true positive predictions among all positive predictions.

### **Class Probability Distribution:**

We visualized the class probability distribution to understand how confident the model is in its predictions. The density plot showed the distribution of predicted probabilities for each class (vaccinated and non-vaccinated).

## **Insights and Conclusion**

The analysis revealed valuable insights into the factors influencing H1N1 vaccine uptake. The Naive Bayes model performed reasonably well, achieving an accuracy of X%. This suggests that the model can effectively predict vaccine uptake based on the selected predictor variables.

Furthermore, the confusion matrix allowed us to understand the trade-offs between different types of prediction errors. The model's high specificity indicates its ability to correctly identify non-vaccinated individuals, which is crucial for public health interventions.

The class probability distribution showed that the model's predictions were well-distributed across different probabilities, indicating that the model's confidence in its predictions was reasonably balanced.

## **Future Directions**

While the Naive Bayes model performed well, further analyses could explore the impact of additional features, conduct cross-validation for more robust results, and compare the Naive Bayes model with other classification algorithms.

In conclusion, this analysis provides insights into the factors influencing H1N1 vaccine uptake and demonstrates the effectiveness of the Naive Bayes classification algorithm in predicting vaccine uptake. These findings can inform public health strategies and interventions.