**Analysis Planning Worksheet**

**Evaluation Question**

How to classify emails as spam or no spam?

**Independent Variable(s)**

These variable(s) are causing something or creating an effect. List what each is and whether it is categorical or continuous. It is ok to only have one.

**Variable**

Text of email

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

**Variable**

Text sequences

□ Categorical : # of levels \_\_\_\_\_ □ Continuous

**Variable**

Model Selection

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

Variable

Implementation

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

Dependent Variable(s)

These variable(s) are influenced by your independent variable and *depend* on them. List what each is and whether it is categorical or continuous. Unless they are related, you should have only one.

**Variable**

Emails

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

**Variable**

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

**Variable**

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

Variable

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

Now that you know the type and number of independent and dependent variables, you are ready to use the analysis flow charts to choose your analysis!

**Analysis:**

To properly classify an email as spam or no spam when it is sent to an individual’s email address based on the text in the email.

**Analysis Planning Worksheet**

**Evaluation Question**

What binary code assists in the determine of classifying an email as spam or no spam?

**Independent Variable(s)**

These variable(s) are causing something or creating an effect. List what each is and whether it is categorical or continuous. It is ok to only have one.

**Variable**

Advertisement Email stating to buy or try something

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

**Variable**

Collections email, you have won

□ Categorical : # of levels \_\_\_\_\_ □ Continuous

**Variable**

Email to get discounts on vacation

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

Variable

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

Dependent Variable(s)

These variable(s) are influenced by your independent variable and *depend* on them. List what each is and whether it is categorical or continuous. Unless they are related, you should have only one.

**Variable**

emails

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

**Variable**

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

**Variable**

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

Variable

□ Categorical: # of levels \_\_\_\_\_ □ Continuous

Now that you know the type and number of independent and dependent variables, you are ready to use the analysis flow charts to choose your analysis!

**Analysis:**

To determine if a code in the email helps to classify it as spam or no spam.