Assessment 2C – Week 4 Workbook Exercise
Logistic Regression

Question	Answer	SAS VA visualisation reference
1. How many rows of data does BIOORGANICS contain?	111,115	
2. How many missing values are there for *Organics Purchase Indicator*?	None	
3. How many customers purchased organic products?	27,525	Cross-tab 1 (not exported)
4. Which gender buys more organic products?	Females	Visualisation 1
5. Which statement most correctly describes the conclusion you can make about those people who buy organic products?	a) People who are younger and more affluent tend to buy more organic products.	Visualisation 2
6. How many observations were used to build the model?	82,040	Logistic 1
7. Which variables are not considered important to the model?	Total Spend, Geographic Region and Neighbourhood Cluster-7 level	Logistic 1 – Fit Summary
8. What is the R square for the new model?	0.2258	Logistic – with IM & VarSel
9. What is the K-S statistic for the new model?	0.4706	Logistic – with IM & VarSel Assessment
10. Which model is selected when the fit statistic is Misclassification?	Logistic – with IM & VarSel	
11. Which model is selected when the prediction cutoff is set to 0.25 and the FDR statistic is chosen?	Logistic 1	