MKTG7510 – Loyalty and Segmentation Project

<YOUR NAME>
<STUDENT NUMBER>

A short introduction to the report can be made here which simply states the contextualisation of your project (e.g., what brand/company/industry it is written for/about) that is no longer than the example text provided here (i.e., one paragraph). For example, "The purpose of this project is to generate consulting advice for how <insert context> can achieve the management objective of developing or maintaining consumer loyalty with marketing communications targeted at consumers attitudes, feelings and perceptions".

Section 1: Data Preparation

Sampling. Data were collected using the online survey platform Qualtrics between <start date> and <end date>. Respondents were strategically targeted in the following ways to maximise sample size. <insert processes>. Table X outlines descriptive statistics about the sample.

Table 1 – Sample

Demographic Traits	
Trait 1	%
Trait 2	n
Trait 3	%
Trait 4	n
Other traits	
Trait 5	%
Trait 6	n
Trait 7	%

Comment on its generalisability of the descriptive statistics in terms of how well you have achieved a sample that is representative of a population of interest.

Factor analysis. Begin with a statement about how whether or not you have reliably measured the variables you intended to measure using your ratings scales. This must be done using factor analysis. Outline what items are included in the creation of composite variables (for example, variable X included items 1, 2, and 3 but excluded item 4 due to crossloading).

Consider including a table of your measures.

Comment on whether the KMO and Bartlett's Test of Sphericity is significant. For example, "A KMO value of .853 was observed which had a significance value of less than 0.05, therefore there are significant patterns of correlation among the survey items". Include the table from SPSS. If on Windows use the snip tool (Windows key + Shift + S).

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.853
Bartlett's Test of Sphericity	Approx. Chi-Square	2140.612
	df	171
	Sig.	.000

Describe the pattern matrix of factor loadings (in terms of whether they conform to your expectations, or whether there are issues like crossloadings/unexpected groupings). Again, include the table from SPSS.

Table 2 – Factor analysis results

Pattern Matrix[a]

	Component			
	1	2	3	4
RQanta1	.893			
RQanta2	.896			
RQanta3	.909			
RQanta4	.975			
RQanta5	.733			
RQanta6	.817			
RQanta7	.634			
RQanta8	.599			
RQanta9				.519
RQanta10				.474
RQanta11			.910	
RQanta12			.914	
RQanta13			.885	
RQanta14		-.945		
RQanta15		-.954		
RQanta16				.939
RQanta17				.933
RQanta18				.985
RQanta19				.819

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

a. Rotation converged in 8 iterations.

There are two options relating to the creation of your composite variables depending on the results of your factor analysis. The first is to be guided by the structure factor loadings in your pattern matrix, which may involve dropping some items or merging variables together if there is a sound theoretical basis to do so.

The second is default back to the measurement model provided by literature from where you originally adopted the rating scales from if your factor analysis results are intractable. In either case, you should describe how—*if at all*—these differ to the original conceptual definitions found in the literature your measures were adopted from.

For example: "In the pattern matrix above, items 1 through 6 group unexpected group together with items 7 and 8. These were meant to measure perceived credibility and perceived quality as two distinct constructs. Given the loadings from the quality items are relatively weak (.634

and .599) and also that there is no theoretical basis to combine these constructs, I have decided default back to the original measurement model provided by Erdem, Swait, Valenzuela (2006)."

Cluster analysis. Identify potential homogenous groups in your sample for segmentation analysis using cluster analysis. You may cluster on any variables in your dataset. When presenting the results of cluster analysis, you need only to report the counts (cluster sizes) and only pick a solution if it has at least 5 members in each cluster.

Show the counts with either a frequency table or bar chart (or both), using the output from SPSS.

Table 3 – Cluster Analysis (5 segments solution)

Average Linkage (Between Groups)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	33	30.6	43.4	43.4
	2	6	5.6	7.9	51.3
	3	20	18.5	26.3	77.6
	4	10	9.3	13.2	90.8
	5	7	6.5	9.2	100.0
	Total	76	70.4	100.0	
Missing	System	32	29.6		
Total		108	100.0		

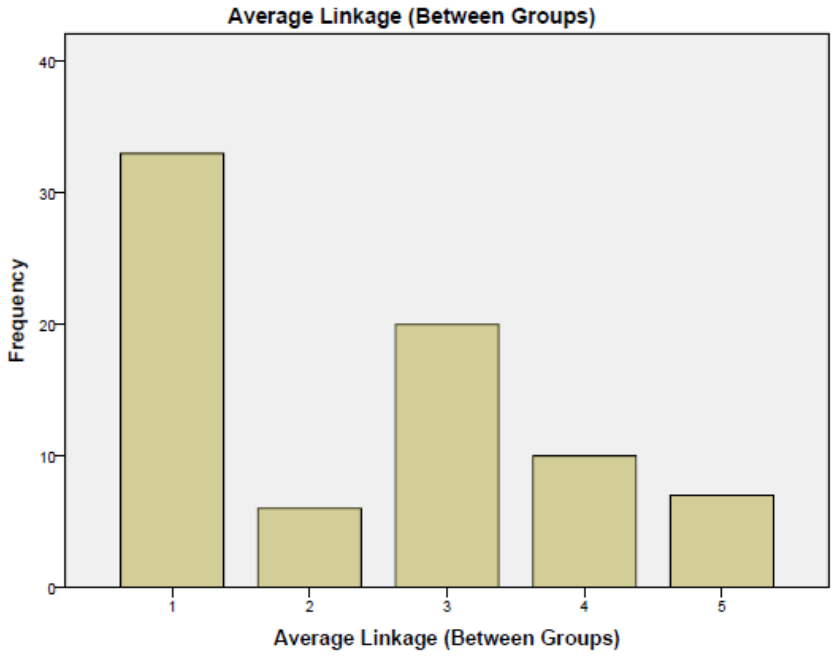


Figure 1 - Cluster Analysis (5 segments solution), bar graph

If no solution has at least 5 members in each cluster, then *you should not select any cluster solution*. If this is the case for your data, explain this in your report. In the next section, you can use some alternative basis for segmentation (e.g., a simple demographic variable or other sample trait).

Section 2: Data Analysis (Loyalty and Segmentation)

The first objective is to provide insights into what might be the potential drivers of loyalty in your chosen industry/context. The main insights should be drawn from a linear regression model (i.e., one predicting loyalty measured as a continuous variable).

The second objective is to provide insights into what might be a useful basis for segmentation in your chosen target population. ANOVA should be used to profile differences between segments.

Additionally, a discriminant analysis model may be used to understand continuous drivers of membership to observed (i.e., non-latent) groups/segments.

Objective 1: Drivers of loyalty

1. **Specify the model(s)**

Within this subsection, list the variables included in your regression model(s), including what are the outcomes of interest and what inputs might explain variance in and predict this outcome. Comment on the appropriateness of these variables for use in your chosen type of regression model(s). For example:

Independent variables:	**Dependent variable**
• Brand Credibility • Perceived Quality • Info Costs Saved • Perceived Risk • Perceived Value	• Loyalty

These variables are appropriate for use in a linear regression model as both the independent variables and dependent variable are measured using continuous rating scales.

A model diagram and/or equation would be useful here (it would communicate the same thing as the above list, but more effectively).

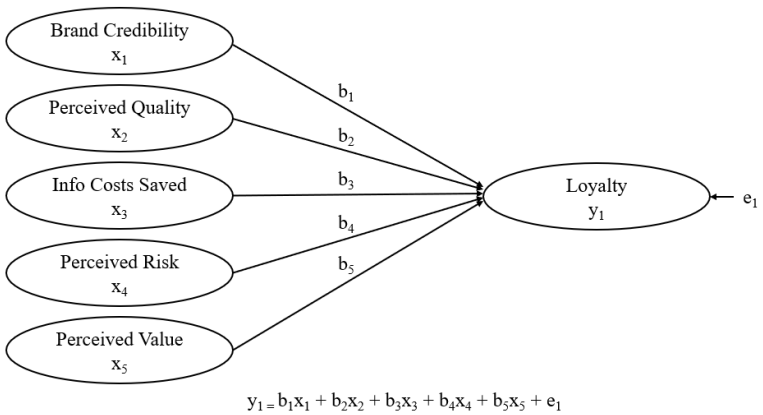


Figure 2 – Regression Model Specification

2. Screen the Input Data

Use frequency analysis, measures of central tendency (mean, standard deviation, minimum and maximum, and skewness and kurtosis).

Statistics

		Q_LOYALTY	Q_CRED	Q_QUAL	Q_ICS	Q_RISK	Q_PRICE
N	Valid	108	108	108	108	108	108
	Missing	0	0	0	0	0	0
Mean		3.5417	3.7901	3.6481	3.4074	2.1821	3.6111
Std. Deviation		1.03134	.76576	.94784	1.01196	1.00918	.88691
Skewness		-.720	-.866	-.345	-.259	.968	-.242
Std. Error of Skewness		.233	.233	.233	.233	.233	.233
Kurtosis		.087	1.414	-.503	-.284	.905	-.562
Std. Error of Kurtosis		.461	.461	.461	.461	.461	.461
Minimum		1.00	1.00	1.00	1.00	1.00	1.50
Maximum		5.00	5.00	5.00	5.00	5.00	5.00

Use visualisations (histograms) to understand how the data are distributed.

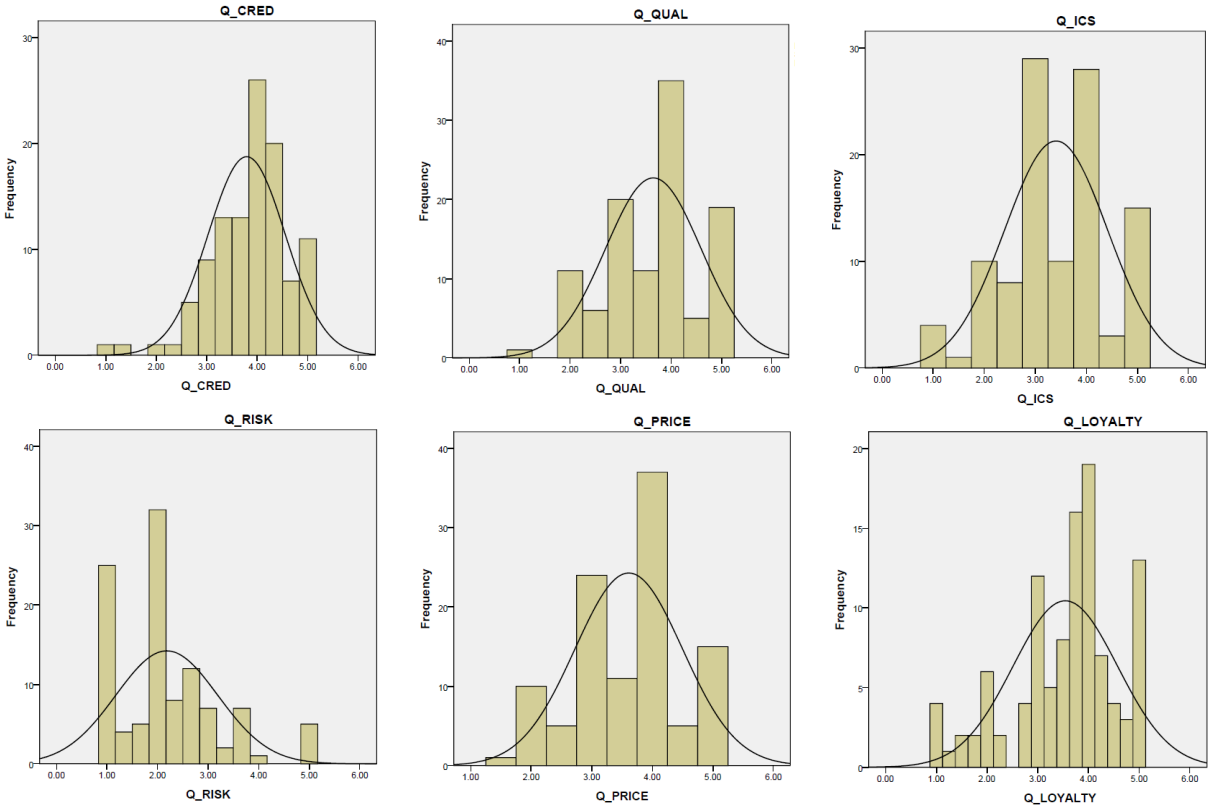


Figure 3 – Panel plot of histograms with normal curves

Use simple scatterplots…

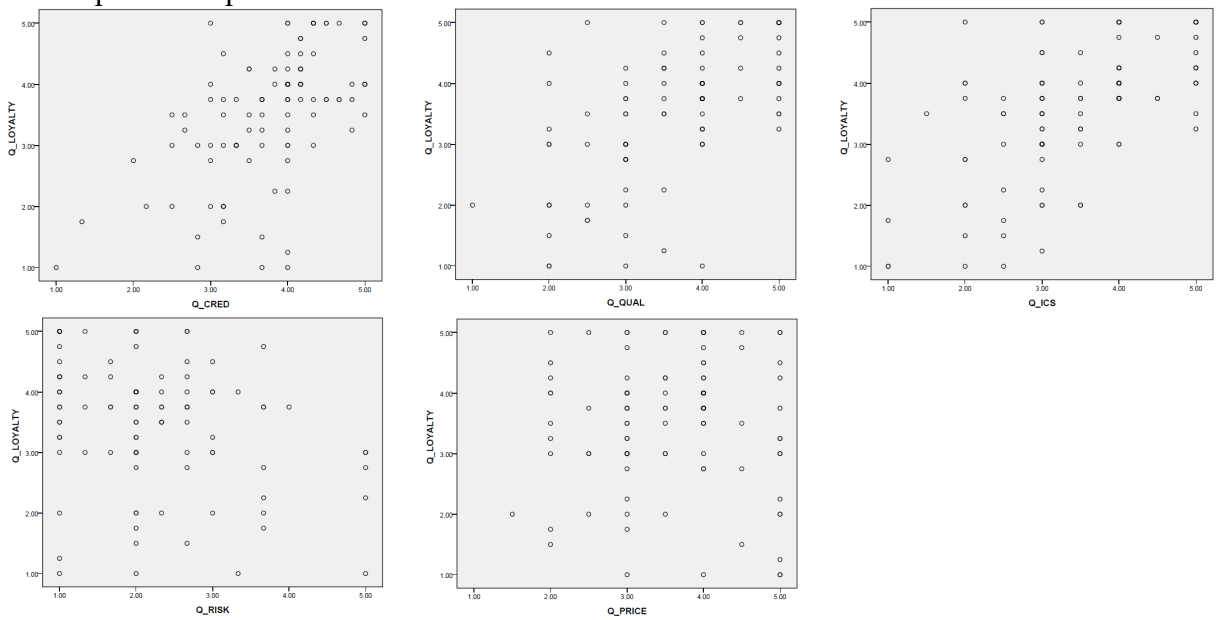


Figure 4 – Panel plot of scatterplots

…and correlation to understand the functional form of the relationships.

Table 4 – Correlation Matrix

Correlations

		Q_LOYALTY	Q_CRED	Q_QUAL	Q_ICS	Q_RISK	Q_PRICE
Q_PI	Pearson Correlation	1	.573**	.609**	.613**	-.322**	.001
	Sig. (2-tailed)		.000	.000	.000	.001	.990
	N	108	108	108	108	108	108
Q_CRED	Pearson Correlation	.573**	1	.753**	.540**	-.364**	.177
	Sig. (2-tailed)	.000		.000	.000	.000	.067
	N	108	108	108	108	108	108
Q_QUAL	Pearson Correlation	.609**	.753**	1	.563**	-.353**	.180
	Sig. (2-tailed)	.000	.000		.000	.000	.062
	N	108	108	108	108	108	108
Q_ICS	Pearson Correlation	.613**	.540**	.563**	1	-.407**	.012
	Sig. (2-tailed)	.000	.000	.000		.000	.905
	N	108	108	108	108	108	108
Q_RISK	Pearson Correlation	-.322**	-.364**	-.353**	-.407**	1	-.221*
	Sig. (2-tailed)	.001	.000	.000	.000		.021
	N	108	108	108	108	108	108
Q_PRICE	Pearson Correlation	.001	.177	.180	.012	-.221*	1
	Sig. (2-tailed)	.990	.067	.062	.905	.021	
	N	108	108	108	108	108	108

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

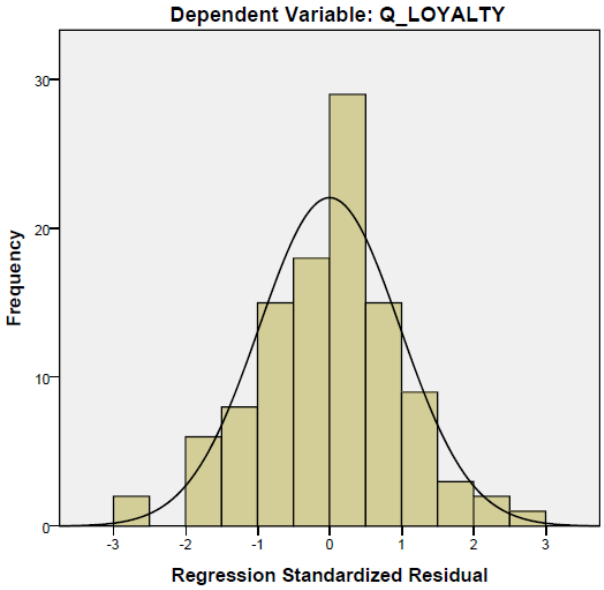
This step provides an initial insight into the possible violation of assumptions.

(Comment on whether there are any possible violations of assumptions based on the analysis outputs you generate for this section).

3. Test assumptions / 4. Estimate the Model

Interpret the output from more robust tests of the assumptions using diagnostics which are produced at the same time as estimation. Check that the following model assumptions have been met using the methods listed in parentheses:

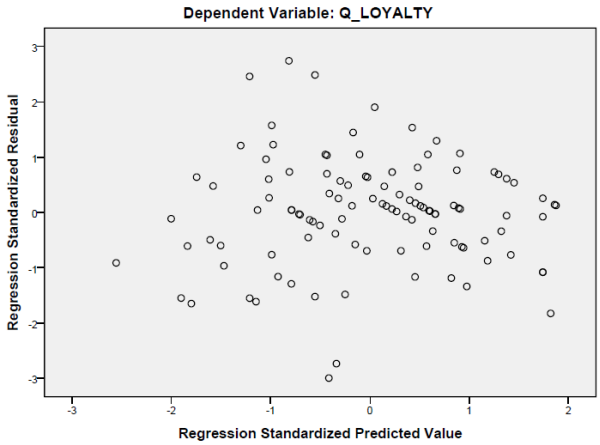
- Normality (look at the distribution of regression residuals),



- Linearity (refer to the simple scatterplots and correlation statistics created before estimation),

See Figure 3 and Table 4. Describe these plots and how they relate to providing evidence about meeting this assumption.

- Homoscedasticity (look at the standardised residual plot, or a Levene's test if using ANOVA, or a Box's M test if using discriminant analysis),



- Independence (refer to your sampling methods).

For each assumption, provide a succinct statement about whether the assumption has been or has been violated. Regardless of whether any or all assumptions are met/violated, you must *proceed with your analysis* to the next step of assessing model fit and interpreting the model output. Satisfying all of the assumptions means the model results will be accurate (although not necessary fitting well or with much significance). Violations of assumptions may lead to biased estimates (over/under estimating the true significance of coefficients).

5. Assess model fit / 6. Interpret the model

Provide an assessment of the amount of variance explained by your model, and whether this amount is statistically significant before interpreting the appropriate model coefficients. The appropriate coefficients depend on the type of model.

You should use a linear regression model to meet the first objective of the report (modelling the drivers of loyalty):

The appropriate coefficients to interpret in the regression model are:

- Adjusted R^2 value
- Significance of the F-test
- Unstandardised coefficients and their significance

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.705[a]	.497	.473	.74884

a. Predictors: (Constant), Q_PVALUE, Q_ICS, Q_PRISK, Q_CRED, Q_QUAL

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	56.615	5	11.323	20.192	.000[b]
	Residual	57.197	102	.561		
	Total	113.813	107			

a. Dependent Variable: Q_LOYALTY

b. Predictors: (Constant), Q_PVALUE, Q_ICS, Q_PRISK, Q_CRED, Q_QUAL

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	.798	.594		1.344	.182
	Q_CRED	.234	.148	.174	1.580	.117
	Q_QUAL	.315	.122	.290	2.592	.011
	Q_ICS	.348	.093	.342	3.752	.000
	Q_PRISK	-.039	.082	-.038	-.474	.637
	Q_PVALUE	-.109	.086	-.094	-1.278	.204

a. Dependent Variable: Q_LOYALTY

Figure 5 – Regression Output (Modelling the Drivers of Loyalty)

Objective 2: Segmentation

If you have a cluster analysis solution with more than 5 individuals per segment, then use this as the "grouping factor" in an ANOVA model. If you do not, then select some other variable (e.g., a demographic trait) for an ANOVA model.

For an ANOVA model you only need to report the difference in means and their significance:

Table 5 – Difference in means of Loyalty across 5 sample clusters

Qloyalty

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	37	3.8649	.94022	.15457	3.5514	4.1783	1.00	5.00
2	34	3.1765	1.07930	.18510	2.7999	3.5531	1.00	5.00
3	21	3.3571	1.13900	.24855	2.8387	3.8756	1.00	5.00
4	7	3.8929	.77536	.29306	3.1758	4.6099	3.00	4.75
5	9	3.7500	.67315	.22438	3.2326	4.2674	2.75	5.00
Total	108	3.5417	1.03134	.09924	3.3449	3.7384	1.00	5.00

Table 6 – Significance of differences means of Loyalty across 5 sample clusters

Qloyalty

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	10.368	4	2.592	2.581	.042
Within Groups	103.444	103	1.004		
Total	113.813	107			

The above compares the level of loyalty between each of the 5 clusters found in the sample which have more than 5 individuals. The difference in mean between these groups is statistically significant, and cluster number 4 has the highest level of loyalty (3.89). Further analysis shows that this cluster is 100% female in its demographic profile (sex = 0, female).

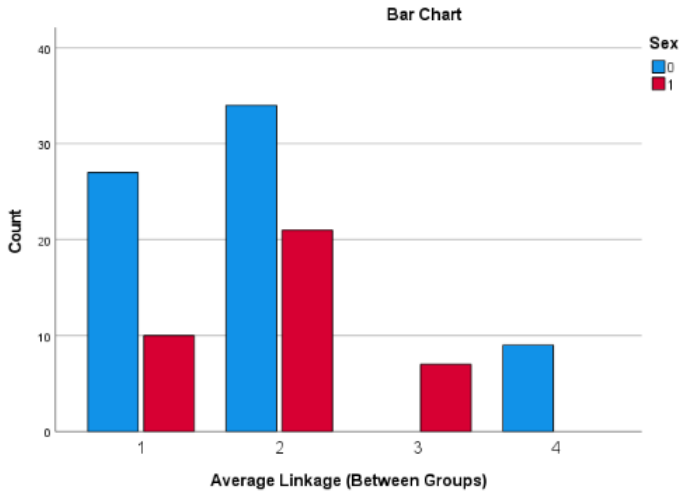


Figure 6 – Demographic Profile of Market Segments

Additional analysis.

If you have conducted other analysis, position it here (after your analysis which answers the two primary objectives of the report). You might, for example, have some observed categorical outcome you wish to analysis using discriminant analysis. If so, then only report the squared canonical correlation and its significance, and/or the hit ratio and the values within the structure matrix. Other additional analysis can also be included here (examples: other types of between or within groups analysis, alternate regression models, analysis of competitors, etc.)

Section 3: Recommendations (1 page)

Objective 1: Drivers of loyalty.

The regression model presented shows that "Information Costs Saved" is the most significant variable in the regression model. This variable was operationalised using the following items "*This airline saves me time shopping around*." and "*This airline saves me time and effort trying to do better*.".

Objective 2: Segmentation.

The cluster analysis and profiling models shows that a segment comprising a majority of females have the highest level of potential loyalty.

Communication Strategy:

The evidence from this data analysis suggests that campaign which focuses on these concepts will be the most effective at increasing consumers loyalty (purchase intentions) for Qantas. Specifically, an advertising campaign should emphasise the time and effort consumers may save by choosing to fly with Qantas and to tailor this message to be appealing to a female demographic. The following billboard mock up was generated by AI using the prompt: *A billboard for Qantas with the message 'Shopping around? Fly Qantas' appealing to women*,

Figure 7 – Example Communication Strategic (Image Created Using AI)

Submission Guidelines

The proposal should be delivered in the form of a professionally formatted written document of no more than **3,000 words (tables/figures excluded)**. The final section (Recommendations) should not exceed 1 page. You may **not** attach any appendices.

You are encouraged to be creative with respect to stylistic elements such as typesetting, spacing, alignment, colours, use of personal brand/agency branding etc. The document should ideally resemble a professionally produced document (it should *not* look a high school essay!).

The proposal will be marked against the rubric (see attached to Turnitin submission link).

Notes about this template

You should follow the layout that this template suggests with respect to the major/minor headings which correspond to the marking rubric. You may add further minor subheadings if this helps communicate your ideas. Note that this template is **not indicative** of recommended word counts in each section (this template has fewer than 1,800 words total). As per above, **you are encouraged to be creative** with respect to other formatting elements. This template has only used a bare minimum of formatting.