GBA424 ANALYTICS DESIGN & APPLICATION

EarlyRider Toy Horse Conjoint Experiment

Presented by **MSBA2 Team#16**:

- Fangyuan Liu (Milar) - 31637503
- Kaili Tan (Kaili) - 31648233
- Xuanhe Xu (Roger) - 31659987
- Ze Long (Kyle) - 31630917

10 Feb 2020

Highlights

REPLACING THE ORIGINAL PRODUCT CHOICE WITH PROFILE 4 AND PROFILE 16.

** Competitor will only reduce his price once; Every consumer will make a purchase; Consumer's taste will remain constant.

Excluding all design within the product itself, consumer **LOVES** product with **LOWER** price.

Consumer with children in different gender has **TOTALLY DIFFERENT** tastes in **motion** and **style**.

Total market share could be **improve** by **90%** with an **increase** in **profit** of **31.4%.**
Product under profile 5 and profile 13 provide a profitability over **70%**.
Current product combination carries the **highest** profitability and ranked as the **3rd** on total profit.

Benefit Segmentation
- A Priori (by Gender)

General preference:

- Low price and large size

Consumer with children gender male:

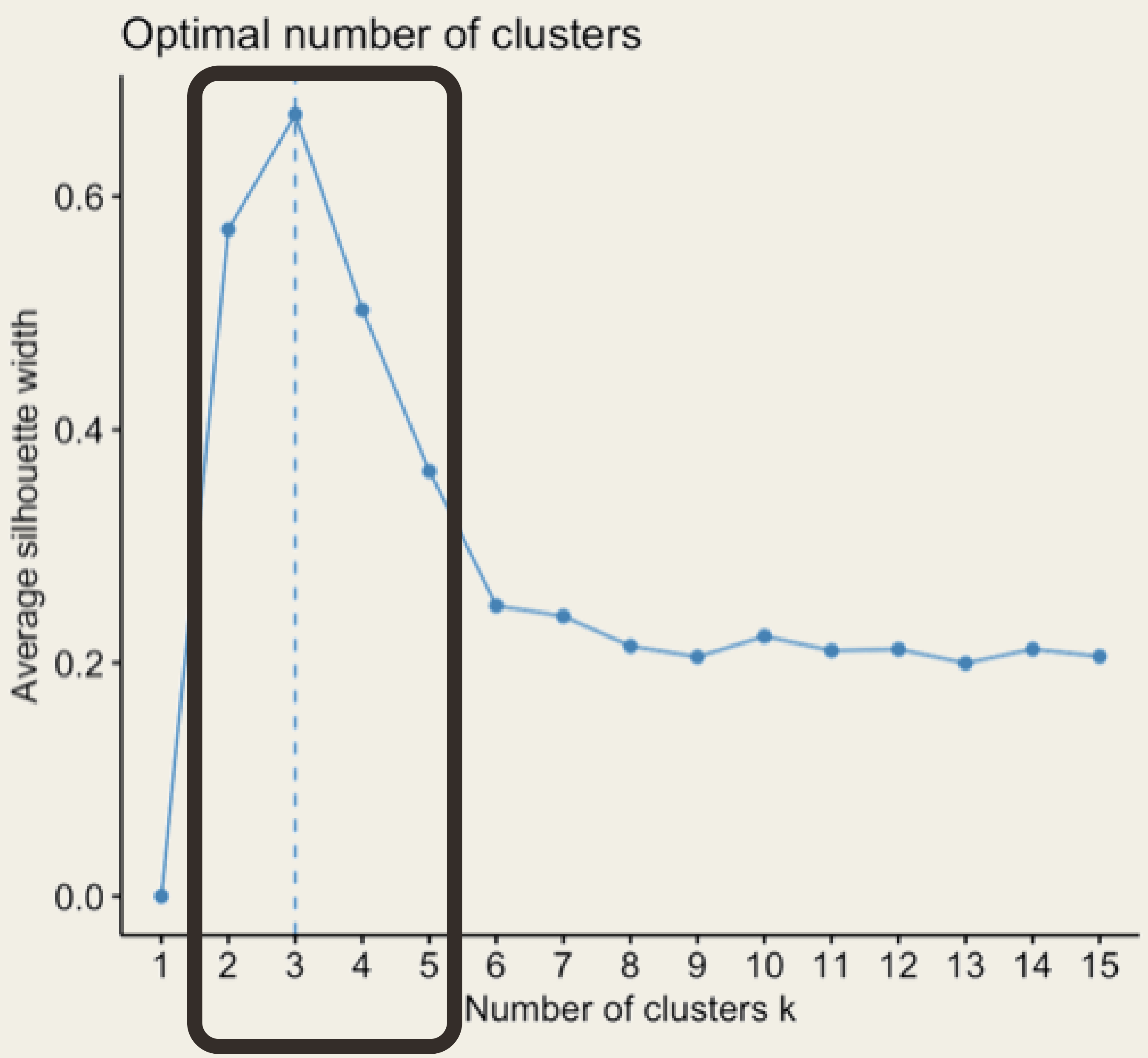
- Bouncing and racing
- Higher baseline on judging products and difficult to satisfy.
- More sensitive to price, the cheaper the higher possibility to make a purchase.

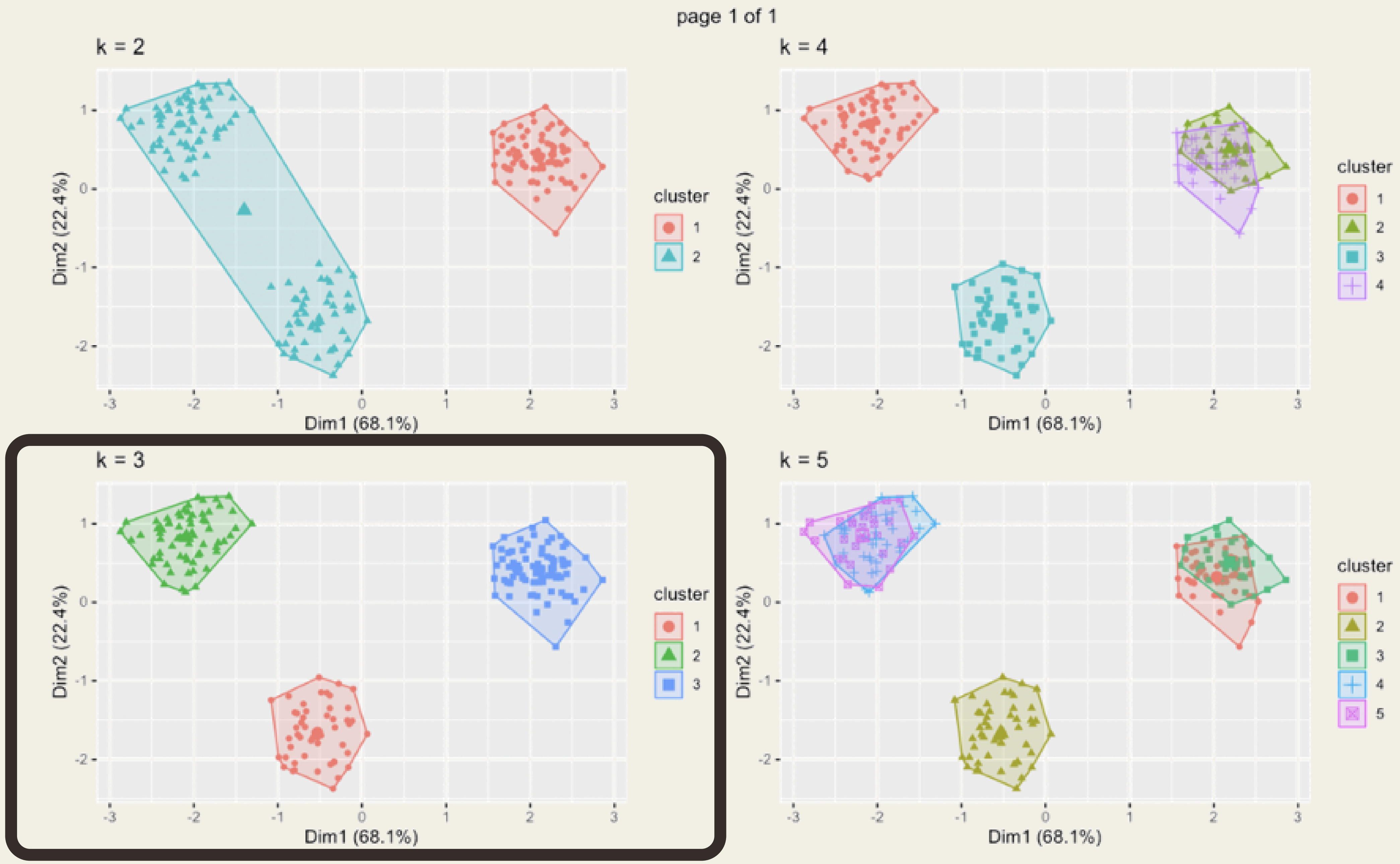
Consumer with children gender female:

- Rocking and glamour
- Lower baseline on judging the product and easier to satisfy by the product.
- Size and style are the 2 major critics while considering the product.

	(Intercept)	price	size	motion	style
male	36.56684	16.85734	3.850932	-0.7601191	-1.889529
female	40.87007	13.50856	7.755489	2.9067692	3.726996

Benefit Segmentation
- Post-Hoc





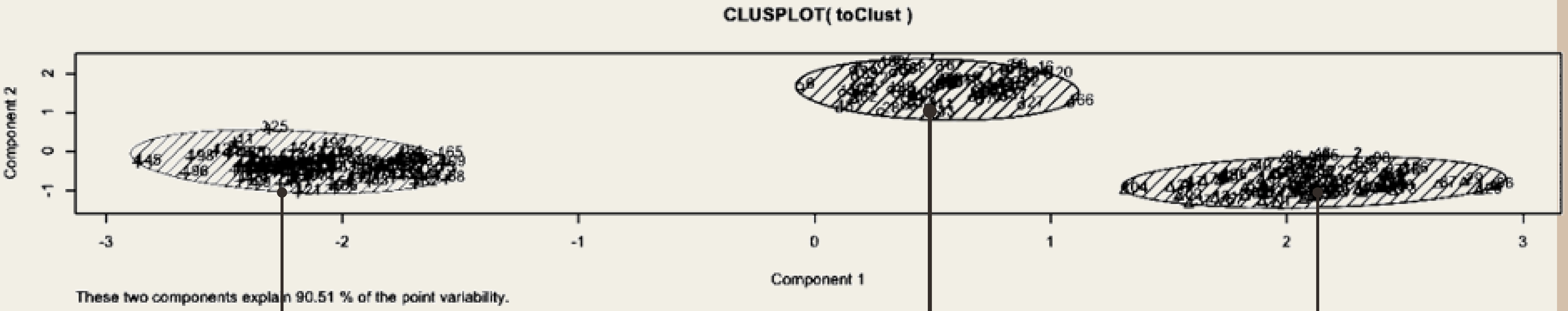
Optimal number of clusters

- elbow rule
- number of cluster to consider: 2, 3, 4, 5

Visualize the potential clustering results -> k = 3

- # of segments > 3: overlap among clusters
- # of segment = 2:
  - samples are not divided informatively and thoroughly

Benefit Segmentation
- Post-Hoc



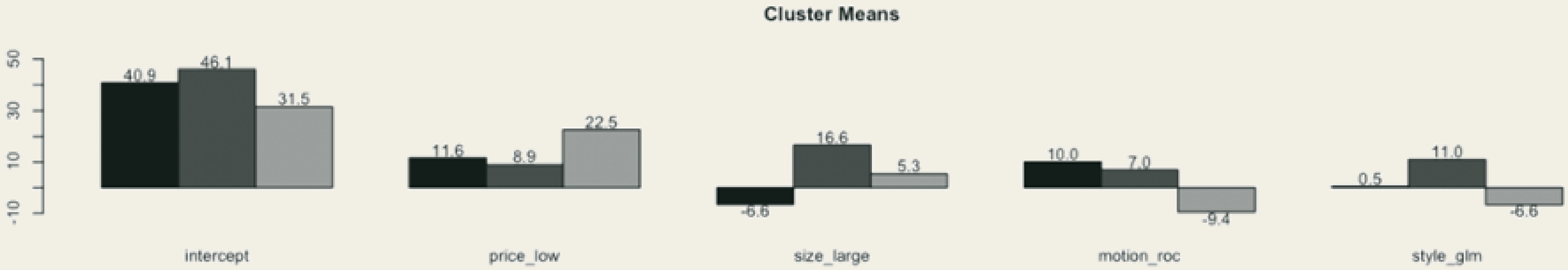
Segment 1:
- prefers rocking and glamour style toy horse
- quite different from the other two

Segment 2:
- prefers small toy horse, which is a unique taste.

Segment 3:
- really loves large size and glamour style toy horse

Benefit Segmentation

- Post-Hoc



Segment 1: profile 4

- Coef(price) = 22.5: price sensitive, always want to buy cheap products.
- Coef(size) = 5.3: don't care much about the size, but larger ones could be better.
- Coefficient (motion) = **-9.4**: **prefer bouncing toy horse which is more sportive**.
- Coefficient (style) = **-6.6**: **prefer racing toy horse. It's also more sportive**. Maybe their kids are lively and active.

Segment 2: profile 14

- Coef(price) = 11.6: like cheaper products.
- Coef(size) = **-6.6**: unlike the other two, **prefer small toys**. Maybe their kids are not old enough to play with large toys.
- Coef(motion) = 10.0: prefer rocking toy horse which is safer.
- Coefficient (style) = 0.5: style has minor effect to them. They don't care about this.

Segment 3: profile 16

- Coef(price) = 8.9: price is not very important but still prefer cheaper products.
- Coef(size) = 16.6: like large toys much more than small ones. Their kids may be 3-4 years old.
- Coef(motion) = 7.0: prefer rocking toy horse, safer than bouncing ones.
- Coef(style) = 11.0: glamour style is much better. Their kids may like good looking toys.

Market Simulation

- Disaggregate choice model using first choice rule

	Profile 1	Profile 2	Profile 3	Profile 4	Profile 5	Profile 6	Profile 7	Profile 8	Profile 9	Profile 10	Profile 11	Profile 12	Profile 13	Profile 14	Profile 15	Profile 16
1	44.63739	45.97442	63.49984	75.38910	43.21365	57.05551	73.57559	78.87591	61.86544	62.58123	71.19551	79.18135	54.195590	74.62188	82.897977	90.20530
2	37.78483	51.08539	29.42126	42.77269	56.30079	69.01173	49.34508	58.01019	40.02285	50.21273	24.35336	42.88283	55.566032	64.67713	36.628834	55.24307
3	51.52104	57.34780	66.05009	73.07050	60.57065	65.25999	71.12056	80.29970	60.17929	66.67391	77.68736	78.36600	61.112546	76.16354	77.886271	86.76004
4	49.66089	56.40732	68.32396	80.84528	52.52966	63.36271	71.83462	83.78224	53.11897	64.11265	78.91761	78.35975	60.421188	71.99519	75.592605	88.12008
5	28.20150	49.72482	33.13775	52.96719	24.92940	45.39599	30.93243	48.84588	19.53881	45.31277	27.09148	57.82775	13.165248	39.82097	13.237254	43.82958
6	41.98542	58.51649	32.09659	50.54725	52.93199	67.94672	31.54656	64.60222	39.98014	58.33498	36.97710	43.28832	44.363571	69.33551	49.175032	59.94890
7	44.43057	52.34401	61.15385	73.24607	50.72389	58.97872	63.98716	75.66603	57.71162	70.81042	79.17099	88.98955	68.686200	74.51352	78.952332	92.18352
8	37.06586	57.90910	41.65336	64.78297	23.85103	45.61303	25.87286	52.31734	27.32691	48.37633	34.02125	48.98299	16.712248	35.83311	21.666269	40.89453
9	29.94362	56.31997	41.00147	66.20057	21.08098	43.01906	27.87135	50.64508	26.53958	48.15846	29.59693	58.87912	12.483027	31.86925	19.067785	43.08548
10	33.94065	51.49321	38.03629	56.04274	23.73545	41.63161	27.04149	46.75015	25.31083	45.48429	28.27256	55.50174	18.360087	30.56294	15.070289	38.19004

Assumption: In consumer model,

- Decisions are rational and based on attribute evaluations, each consumer purchase one product
    - each consumer chooses the profile with his/her highest rating

Rating Data -> Test Values (0/1)

- highest rating: 1 (to purchase), others: 0 (not to purchase)

Market Simulation

- Develop 10 scenarios and calculate market shares

	Profile 4	Profile 5	Profile 7	Profile 8	Profile 13	Profile 14	Profile 16
1	NA	0.220	0.57	NA	0.210	NA	NA
2	0.415	0.040	NA	0.545	NA	NA	NA
3	NA	0.005	NA	0.700	NA	0.295	NA
4	NA	0.020	NA	0.500	NA	NA	0.480
5	0.415	NA	NA	0.520	0.065	NA	NA
6	NA	NA	NA	0.700	0.005	0.295	NA
7	NA	NA	NA	0.495	0.035	NA	0.470
8	0.400	NA	NA	0.335	NA	0.265	NA
9	0.355	NA	NA	0.180	NA	NA	0.465
10	NA	NA	NA	0.405	NA	0.230	0.365

Findings:

- Compared with original scenario, our best scenario obtains 39% market share from competitor.
- With lower price products, to gain market share is to introduce product that satisfies wide range of consumers.

Assumption:

- If we launch the product with lower price, the competitor will respond by lowering price to 119.99 (without changing their wholesale price).
  - switching over from profile 7 to profile 8 -> we earn less revenue when considering **competitive impact**

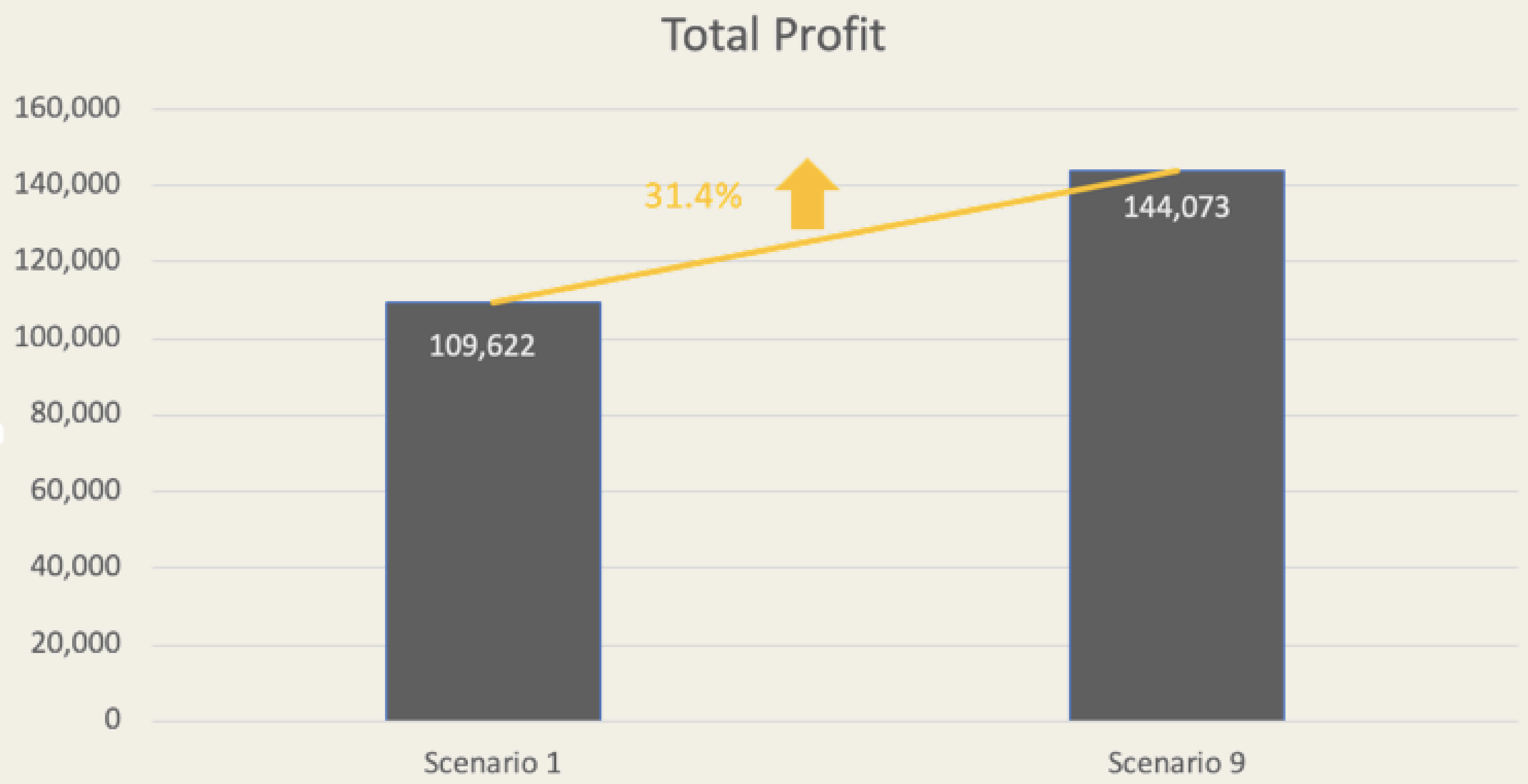
Result: using market share as the criteria,

- best scenario: withdrawing profile 5 and 13, launching profile **4 and 16**
  - profile 4: bouncing and racing (for male kids), profile 16: rocking and glamour (for female kids)

Market Simulation

- Profit for each scenario

	Profile 4	Profile 5	Profile 7	Profile 8	Profile 13	Profile 14	Profile 16	profit
9	0.355	0.000	0.00	0.180	0.000	0.000	0.465	144073.87
8	0.400	0.000	0.00	0.335	0.000	0.265	0.000	120620.07
1	0.000	0.220	0.57	0.000	0.210	0.000	0.000	109622.80
5	0.415	0.000	0.00	0.520	0.065	0.000	0.000	87154.13
10	0.000	0.000	0.00	0.405	0.000	0.230	0.365	84902.87
2	0.415	0.040	0.00	0.545	0.000	0.000	0.000	78455.13
7	0.000	0.000	0.00	0.495	0.035	0.000	0.470	68893.13
4	0.000	0.020	0.00	0.500	0.000	0.000	0.480	65873.33
3	0.000	0.005	0.00	0.700	0.000	0.295	0.000	29401.33
6	0.000	0.000	0.00	0.700	0.005	0.295	0.000	29401.33



Best scenario:

- replacing profile 5 and 13 with profile 4 and 16
  - increase profit by approx. $40,000, compared to original scenario
  - occupy 82% of the market

Market Simulation

- Profitability of each product and the firm overall

Profile 4	Profile 5	Profile 13	Profile 14	Profile 16
0.000	0.220	0.210	0.000	0.000
0.415	0.040	0.000	0.000	0.000
0.000	0.005	0.000	0.295	0.000
0.000	0.020	0.000	0.000	0.480
0.415	0.000	0.065	0.000	0.000
0.000	0.000	0.005	0.295	0.000
0.000	0.000	0.035	0.000	0.470
0.400	0.000	0.000	0.265	0.000
0.355	0.000	0.000	0.000	0.465
0.000	0.000	0.000	0.230	0.365

Market Share

Profile 4	Profile 5	Profile 13	Profile 14	Profile 16	Total profitability	Profit
0.000	0.536	0.527	0.000	0.000	0.7326611	109622.80
0.531	-0.317	0.000	0.000	0.000	0.6270301	78455.13
0.000	-7.609	0.000	0.421	0.000	0.3865138	29401.33
0.000	-1.359	0.000	0.000	0.428	0.5853326	65873.33
0.531	0.000	0.084	0.000	0.000	0.6512749	87154.13
0.000	0.000	-7.609	0.421	0.000	0.3865138	29401.33
0.000	0.000	-0.466	0.000	0.425	0.5961687	68893.13
0.524	0.000	0.000	0.394	0.000	0.6934045	120620.07
0.502	0.000	0.000	0.000	0.424	0.7298309	144073.87
0.000	0.000	0.000	0.354	0.383	0.6141869	84902.87

Profitability

Findings:

- Low market share often results in low profitability, in this case.
- Better not launch a low-priced profile when there is high-priced similar profile in the market.
    - concern: cannibalization of high-priced product
    - high-high or low-low
- High profit does not necessarily mean high profitability.
    - considering extra-costs, such as switching cost (in short-term)

EARLYRIDER

A.1

Appendix Complete Ratings

ID	profile	ratings	price	size	motion	style
1	1	44.63739	0	0	0	0
1	2	45.97442	1	0	0	0
1	3	63.49984	0	1	0	0
1	4	75.38910	1	1	0	0
1	5	43.21365	0	0	1	0
1	6	57.05551	1	0	1	0
1	7	73.57559	0	1	1	0
1	8	78.87591	1	1	1	0
1	9	61.86544	0	0	0	1
1	10	62.58123	1	0	0	1
1	11	71.19551	0	1	0	1
1	12	79.18135	1	1	0	1
1	13	54.19559	0	0	1	1
1	14	74.62188	1	0	1	1
1	15	82.89798	0	1	1	1
1	16	90.20530	1	1	1	1

- Step 1:
  - for each respondent, run regression: ratings by four key attributes
- Step 2:
  - predict ratings for missing profiles (3, 6, 10, 16) using their attributes info
- Step 3:
  - combine the existing profile ratings and those predicted profile ratings in right order

A.2

```
Call:
lm(formula = ratings ~ price + size + motion + style, data = cj.comb)

Residuals:
    Min      1Q  Median      3Q     Max 
-41.254 -12.127  -2.063  11.843  44.221 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  38.8906     0.6291  61.818   <2e-16 ***
price        15.0490     0.5627  26.744   <2e-16 ***
size          5.9594     0.5627  10.591   <2e-16 ***
motion        1.2200     0.5627   2.168   0.0302 *  
style         1.1434     0.5627   2.032   0.0422 *  
```

Regression 1: attributes, without segmentation

With counting survey data as a general population, the consumer are more sensitive to price and size of the product. Consumers prefer product with lower price and larger size. Consumers tend not interested in the motion and style of the product.

EARLY RIDER

Appendix A Priori

Build regression models to conduct a priori segmentation at segment level.

- two segment levels:
  - age, gender

Appendix A Priori

A.3

Regression 2:
attributes, segmented by gender

When taking gender into account, consumer with girls are more sensitive to price compared to those who had boys and they will be less concern about size, motion and style while making a purchase compare to those consumer with boys.

```
Call:
lm(formula = ratings ~ price * gender + size * gender + motion *
    gender + style * gender, data = cj.comb)

Residuals:
   Min     1Q Median     3Q    Max
-45.67 -10.97  -0.52  10.74  47.02

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    36.5668     0.8801  41.547  < 2e-16 ***
price          16.8573     0.7872  21.414  < 2e-16 ***
gender          4.3032     1.1977   3.593 0.000332 ***
size            3.8509     0.7872   4.892 1.05e-06 ***
motion         -0.7601     0.7872  -0.966 0.334327
style          -1.8895     0.7872  -2.400 0.016440 *
price:gender   -3.3488     1.0713  -3.126 0.001788 **
gender:size     3.9046     1.0713   3.645 0.000272 ***
gender:motion   3.6669     1.0713   3.423 0.000627 ***
gender:style    5.6165     1.0713   5.243 1.68e-07 ***
```

Regression 3:
attributes, segmented by age

By adding interaction variables into the regression, consumer with kids that are in the range of 3-4 year will more concern about the products motion and size compared to those having younger kids. Consumer will be more likely to buy a larger size toy horse and more prefer having a bouncing horse.

```
Call:
lm(formula = ratings ~ price * age + size * age + motion * age +
    style * age, data = cj.comb)

Residuals:
    Min      1Q  Median      3Q     Max
-43.528 -11.829  -2.007  11.634  44.386

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  39.5462     0.8917  44.350  < 2e-16 ***
price        14.4133     0.7975  18.072  < 2e-16 ***
age          -1.2982     1.2548  -1.035 0.300928
size          3.8532     0.7975   4.831 1.42e-06 ***
motion        2.7950     0.7975   3.504 0.000464 ***
style         1.1867     0.7975   1.488 0.136877
price:age     1.2588     1.1223   1.122 0.262095
age:size      4.1708     1.1223   3.716 0.000206 ***
age:motion   -3.1188     1.1223  -2.779 0.005486 **
age:style    -0.0857     1.1223  -0.076 0.939140
```

```
Call:
lm(formula = ratings ~ price * gender * age + size * gender *
    age + motion * gender * age + style * gender * age, data = cj.comb)

Residuals:
    Min      1Q  Median      3Q     Max
-46.179 -10.319  -0.684  10.320  49.241

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        37.7780     1.1634  32.473  < 2e-16 ***
price              15.3906     1.0405  14.791  < 2e-16 ***
gender              3.7244     1.6884   2.206 0.027467 *
age                -2.7857     1.7643  -1.579 0.114457
size                2.3625     1.0405   2.270 0.023250 *
motion              1.6932     1.0405   1.627 0.103789
style              -0.9823     1.0405  -0.944 0.345214
price:gender       -2.0586     1.5102  -1.363 0.172928
price:age           3.3735     1.5781   2.138 0.032616 *
gender:age          1.6661     2.4008   0.694 0.487749
gender:size         3.1400     1.5102   2.079 0.037679 *
age:size            3.4235     1.5781   2.169 0.030125 *
gender:motion       2.3208     1.5102   1.537 0.124452
age:motion         -5.6426     1.5781  -3.576 0.000355 ***
gender:style        4.5688     1.5102   3.025 0.002504 **
age:style          -2.0865     1.5781  -1.322 0.186193
price:gender:age   -3.0608     2.1474  -1.425 0.154144
gender:age:size     0.5655     2.1474   0.263 0.792301
gender:age:motion   3.6823     2.1474   1.715 0.086482 .
gender:age:style    2.3354     2.1474   1.088 0.276866
```

Appendix A Priori

A.4

Regression 4: attributes, segmented by both gender and age:

When taking all possible variables into account, the dataset has been divided into minimal parts that only price, affects the general population. Children's age, and gender become factors that affects consumer making decisions while considering motion and style of the product.

Appendix A Priori

A.5

EARLY RIDER

It seems that regression 2 (attributes, segmented by gender) is more significant than other 3 regressions. Then we did a segment – level analysis (segmented by gender). We ran regression for each segment separately. As indicated by the above results, parents of boys and girls both prefer lower price and large size. Parents of boys tend to buy them bouncing motion and racing style toy horse, while parents of girls tend to buy them rocking motion and glamour style toy horse.

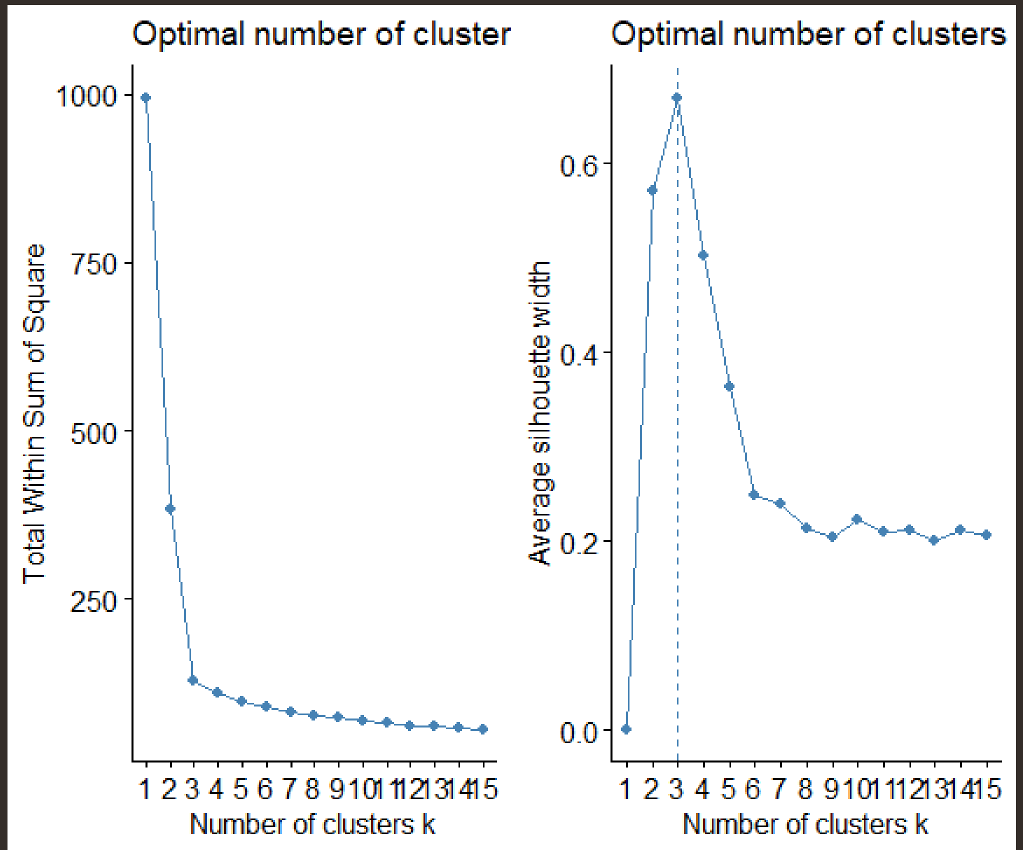
```
Call:
lm(formula = ratings ~ price + size + motion + style, data = cj.m)

Residuals:
    Min      1Q  Median      3Q     Max 
-33.917  -9.740  -2.504   7.017  47.017 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  36.5668     0.7991  45.760  < 2e-16 ***
price        16.8573     0.7147  23.585  < 2e-16 ***
size          3.8509     0.7147   5.388 8.29e-08 ***
motion       -0.7601     0.7147  -1.063  0.28773    
style        -1.8895     0.7147  -2.644  0.00829 ** 
```

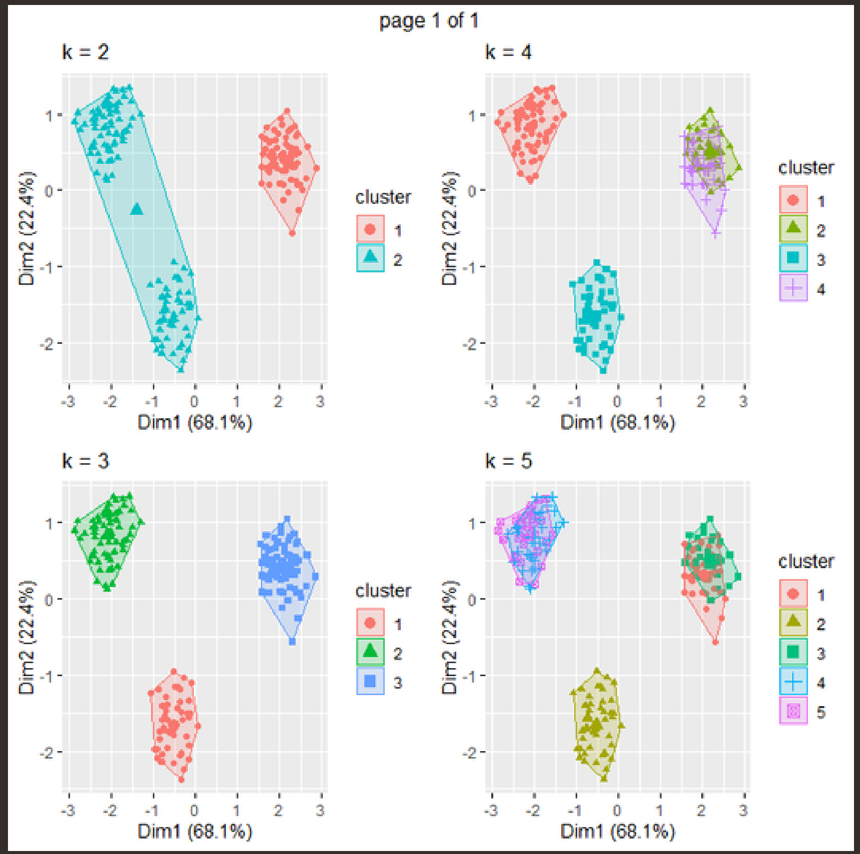
```
Call:
lm(formula = ratings ~ price + size + motion + style, data = cj.f)

Residuals:
    Min      1Q  Median      3Q     Max 
-45.675 -12.784   2.696  12.527  36.175 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   40.870      0.871  46.926  < 2e-16 ***
price         13.509      0.779  17.341  < 2e-16 ***
size           7.755      0.779   9.956  < 2e-16 ***
motion         2.907      0.779   3.731 0.000197 ***
style          3.727      0.779   4.784 1.86e-06 ***
```

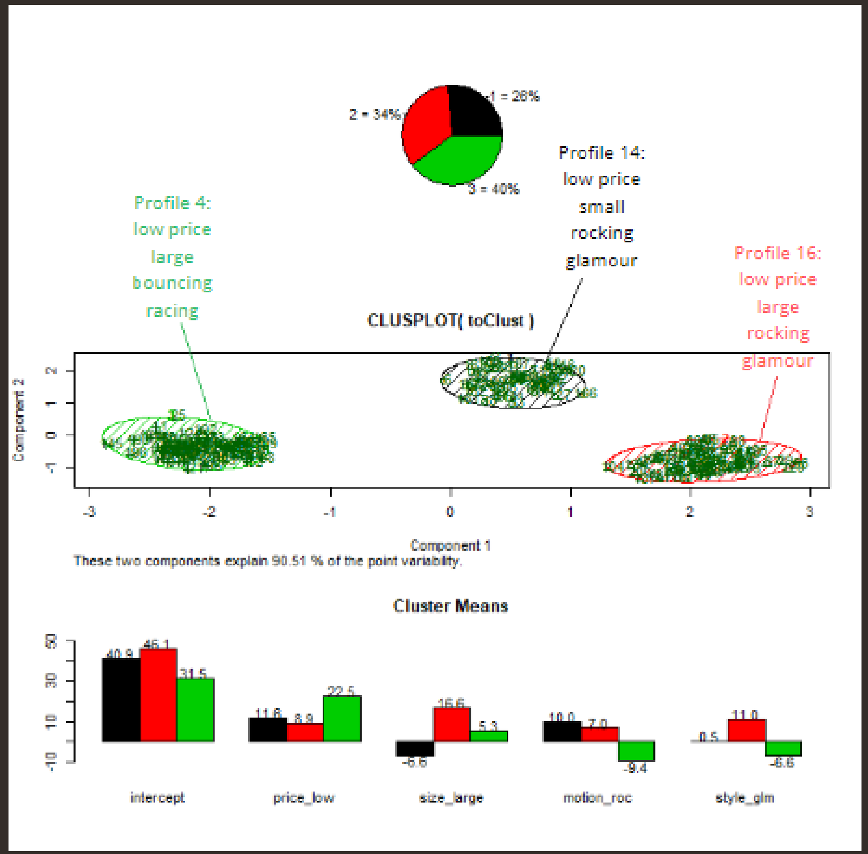
	(Intercept)	price	size	motion	style
male	36.56684	16.85734	3.850932	-0.7601191	-1.889529
female	40.87007	13.50856	7.755489	2.9067692	3.726996

Appendix
Post-Hoc

A.6





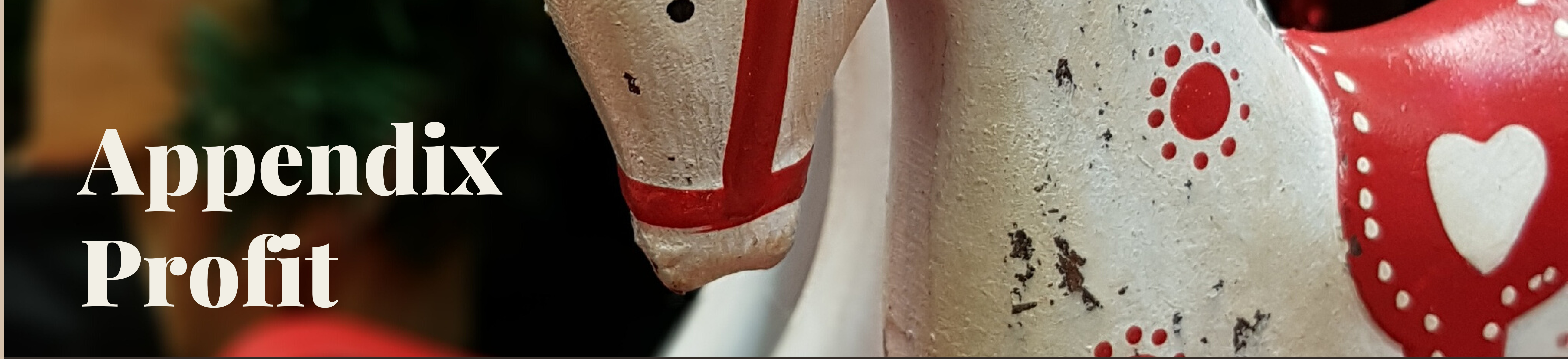


We evaluated number of clusters to use on data with visualizations. The results were a list of weighted sum of squares and the pamk output including optimal number of clusters (nc)
The graphic showed us that dividing the sample into 3 clusters is the best choice.

To confirm that, we also looked at the cluster plots. In the previous graphic, k = 2, 3, 4, 5 are the 4 highest points so we had these 4 plots. By comparing them, we confirmed that we should divide the sample into 3 clusters.

Lastly, we looked at the 3 clusters in more depth. The result shows us that 26% of the respondents prefer small, rocking and glamour horse with low price. 34% of the respondents prefer large, rocking and glamour horse with low price. 40% of the respondents prefer large, bouncing and racing horse with low price.

A.7



Appendix Profit

Profit computation: Assuming all target consumer in the market will make a purchase.
The total marginal revenue is computed by sales price minus the cost of each product and times the quantity of product sold in each profile. Profit = marginal revenue – total cost.
Profitability computation: Marginal revenue of each product/product line divided by the total revenue of each product/product line.

Profile 4	Profile 5	Profile 13	Profile 14	Profile 16	Total profitability	Profit
0.000	0.536	0.527	0.000	0.000	0.7326611	109622.80
0.531	-0.317	0.000	0.000	0.000	0.6270301	78455.13
0.000	-7.609	0.000	0.421	0.000	0.3865138	29401.33
0.000	-1.359	0.000	0.000	0.428	0.5853326	65873.33
0.531	0.000	0.084	0.000	0.000	0.6512749	87154.13
0.000	0.000	-7.609	0.421	0.000	0.3865138	29401.33
0.000	0.000	-0.466	0.000	0.425	0.5961687	68893.13
0.524	0.000	0.000	0.394	0.000	0.6934045	120620.07
0.502	0.000	0.000	0.000	0.424	0.7298309	144073.87
0.000	0.000	0.000	0.354	0.383	0.6141869	84902.87