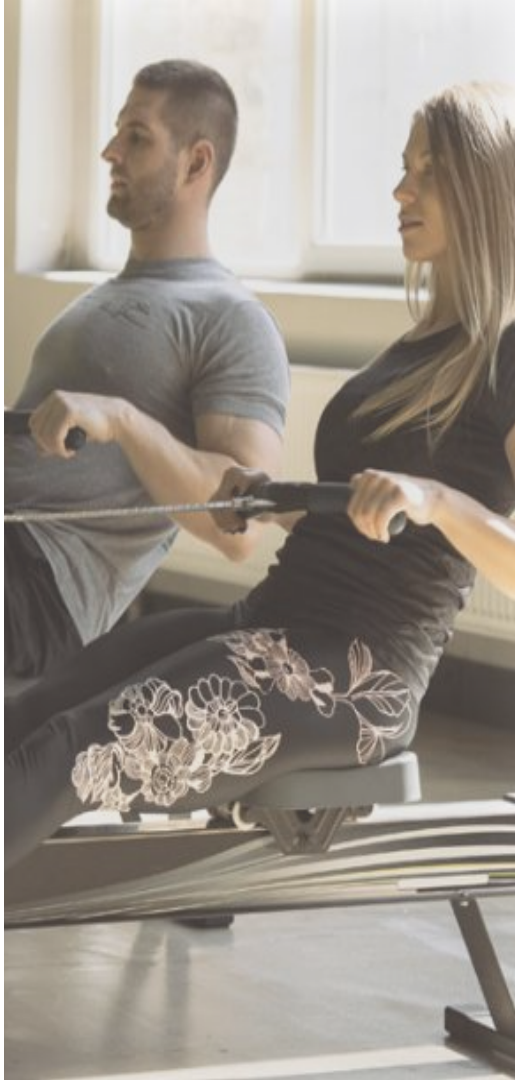
# Strategies for Crew's Cup Fitness

MKT 512 IA2 | Sihan Chen

# Table of Contents

# 01 Recommendations

1. **Faithful segment should be Crew's Cup's primary target.**
    a. This segment has the lowest intent to purchase, which indicates they are very likely to churn compared with the other 2 segments.
    b. Even though this segment has the lowest population, they take the most number of classes, generate most profits, and have stayed the longest with us. Thus they are valuable.
2. **Launch the campaign 'Beneath the Surface' to signal the Faithful customers that it is getting back to its fundamental values to reduce churn.**
    a. The Faithful has low normative commitment, which also has a large influence on intent.
    b. The campaign will also resonate with the new Fitness Buffs who are drawn by the brand's values (segment 3). This is the largest segment and has a large potential for profitability with highest intent.
3. **Crew's Cup can introduce a new type of class to provide long-term customers a new experience that might renew their interest in the brand, if the budget is enough.**
    a. The Faithful has highest forced commitment and forced commitment is proved to have the largest influence on intent to purchase; introducing new classes can renew the interest.
    b. However, this might not have strong effect on the other 2 segments because segment 3 are new customers and has strong interest, and segment 1 only cares about price.
4. **Although the Budget Hunters have a large population and they do generate relatively high profits, Crew's Cup might not consider creating the subscription program.**
    a. Economic and habitual commitment have very small influences on intent to purchase and they are not significant, thus the program might not reduce churn substantially.
    b. The Budget Hunters have high level of forced commitment and they have high intent to purchase, thus they are not likely to churn



# 02 Background

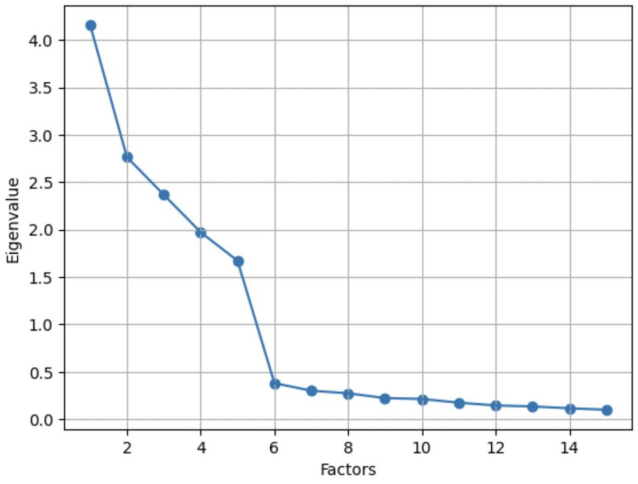
Crew's Cup Fitness, a once successful fitness brand that faced a decline due to management changes, operating decisions, and the COVID-19 pandemic. To **reduce churn**, the management is considering three options: a brand campaign, a subscription-based pricing model, or introducing a new class.

In the following analyses, factor analysis, cluster analysis, regression, and other statistical techniques will be leveraged to:

- test the 5 factor commitment model
- identify and label customer segments and find commitment level differences across segments
- identify significant commitments that have high influence on intent to purchase
- develop future strategies to reduce churn

# 03 Factor Analysis

**The factor analysis validates that the descriptors can be grouped into 5 factors**, given that the eigenvalue > 1 at 5 factors level (Appendix II). Also, there's no significant correlation between factor groups, which means the factor analysis is good (Appendix III). In the factor loading plot, there are 3 variables that have significantly high loadings for each factor (shown as light yellow), for example, f1, f2, and f3 have a significantly larger loading score for Factor 1. This helps us conclude that the commitment model consist of the 5 factors, which are forced committment, habitual committment, economic commitment, normative commitment, affective commitment, and normative commitment.




Scree Plot


Factor Loadings Matrix

Factor 1: f1, f2, f3 — forced commitment
Factor 2: h1, h2, h3 — habitual commitment
Factor 3: e1, e2, e3 — economic commitment
Factor 4: a1, a2, a3 — affective commitment
Factor 5: n1, n2, n3 — normative commitment

### KMO for Factorability

The KMO analysis is runned to measure if the sampling adequacy is a better measure of factorability. With a KMO score of 0.758 (Appendix I), the factor analysis is good to run.

# 03 Factor Analysis - reliability

**Overall, the factor analysis is reliable.** Chronbach's Alpha score is calculated for each commitment to assess the reliability of the factor analysis. All the commitments or factors has a alpha score > 0.8 (Appendix III), which is excellent and indicates there is reliability. The picture on the right also shows how the score changes when one element in a factor is drop. For example, the score for forced commitment overall is 0.932, and after dropping f1, the score becomes 0.913, which means reliability decreases and f1 should be added in the factor. The results indicate that dropping every single variable will result in a decrease in score of the corresponding factor, which means all of the variables are important to include in each factor, and overall the factor model is reliable.


```
{'Forced Commitment': (0.932193670871478,
   {'f1': 0.9134541449088769,
    'f2': 0.8592768355052149,
    'f3': 0.9294605931461686}),
 'Habitual Commitment': (0.931917761455404,
   {'h1': 0.9154531682884581,
    'h2': 0.8683826670963068,
    'h3': 0.9209912003428499}),
 'Economic Commitment': (0.9210877136935092,
    'e1': 0.9133351823847611,
    'e2': 0.8420058269708965,
    'e3': 0.8974122855783895}),
 'Affective Commitment': (0.89553060885824,
    'a1': 0.8907503678217545,
    'a2': 0.7810166326692314,
    'a3': 0.8750356376429118}),
 'Normative Commitment': (0.8864822417434761,
    'n1': 0.8809586277638468,
    'n2': 0.7656999617432587,
    'n3': 0.8580397088084861})}
```


| | Affective | Normative | Economic | Forced | Habitual |
|---|---|---|---|---|---|
| Affective | 1.00 | | | | |
| Normative | 0.01 | 1.00 | | | |
| Economic | -0.13 | -0.22 | 1.00 | | |
| Forced | -0.01 | -0.23 | 0.06 | 1.00 | |
| Habitual | -0.06 | -0.16 | 0.33 | 0.12 | 1.00 |

This is the factor correlation plot, showing the correlation between each factor. There's no significant correlation between factor groups, which means that all variables in each factor are distinct from others in other factors, and the factor analysis is good and reliable.

Given that the the descriptors can be grouped into 5 factors and the factor model is reliable, **Crew's Cup can rely on the factor model to calculate commitment levels for different market segments, thus understanding each segment and developing tailored retention strategies for targets.**

# 04 Cluster Analysis

**There are 3 customer segments.** By running the cluster analysis, there is a big shift at 3 clusters as shown in the scree plot. The scree plot compares the sum of squared error (SSE) for each cluster solution. A good cluster solution might be when the SSE slows dramatically, creating an 'elbow'.

The 1500 survey respondents are grouped into 3 segments. Segment 2 has the smallest population, and the number of people in segment 1 and 3 are roughly the same.

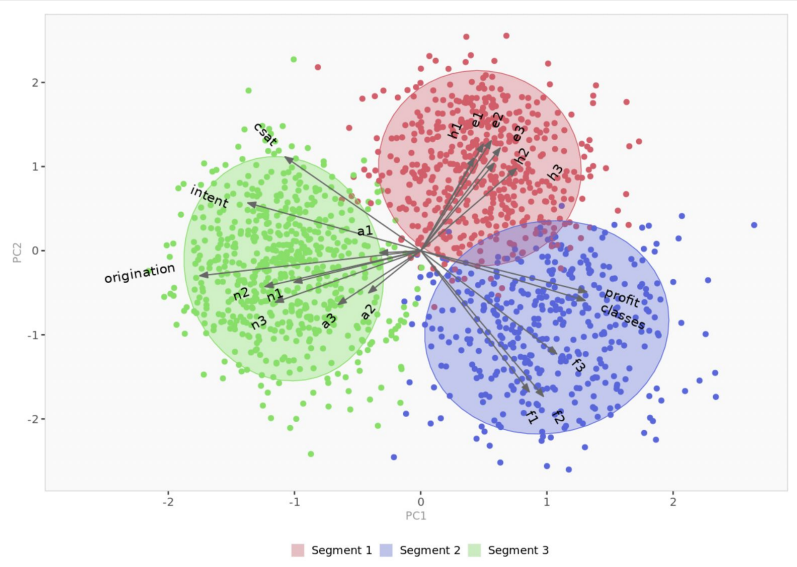



| | Population | Segment 1 | Segment 2 | Segment 3 |
|---|---|---|---|---|
| **Size** | 1 500 | 513 | 392 | 595 |
| **Relative size** | 100% | 34% | 26% | 40% |

Segment Size

# 05 Segments

Based on segment descriptions from  k-means clustering (Appendix IV), the results revealed more information about the segments proposed by the agency. For example, the result shows that there is a group of people who are emotionally attached to the brand, which the agency doesn't reveal. The data is scaled by column and the dimension is reduced with PCA, thus the model is appropriate.




**Segment Space** is plotted after PCA analysis

1

**Budget Hunter & Adjacents.** People in cluster 1 go to Crew's Cup as a habit and are attracted because of the economic benefits. These people might live very close to a Crew's Cup store and they go there for convenience. They love fitness and need exercising regularly so they need a place that's not far from their home but also cheap. There intent and satisfaction are relatively high. This cluster seems to be a combination of the Budget Hunter and the Adjacents proposed by the agency.

2

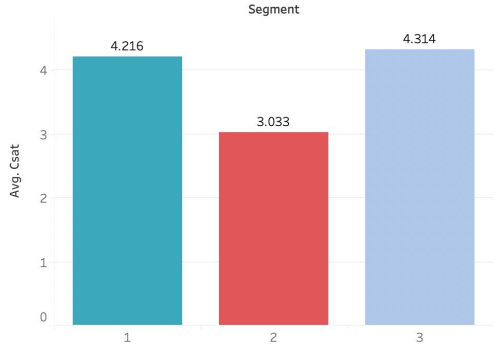
**The Faithful.** Cluster 2 are people who have attended the most classes and are the most profitable. They are also the 'oldest' customers. However, the reason they go there so frequently is that they have no other choices. They might tried other options but ultimately return because there's no better options. Their intent to purchase and satisfaction is the lowest. This matches with the agency's Faithful segment.

3

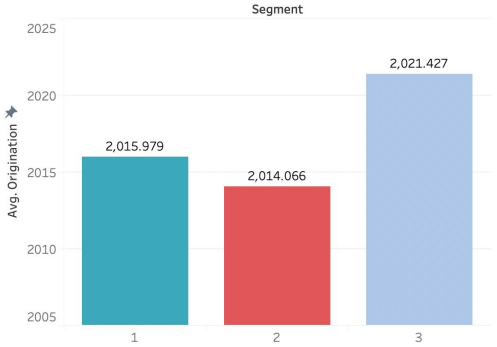
**New Fitness Buffs drawn by values.** People in cluster 3 feel a deep emotional connection and affection to Crew's Cup based on the resonance of their personal values with those upheld by the company. They have highest satisfaction and intent to purchase, and they are new to the brand. They are probably rich due to low economic commitment. This cluster is very similar to the Fitness Buffs described by the agency, who engages in many fitness activities and Crew's Cup is a fraction of their time (because the number of classes is low) and they viewed the price as cheaper than many gym memberships. What's new is that this group of people is also drawn by the values of Crew's Cup and has a strong affection for it.

# 06 Target Segment

Csat

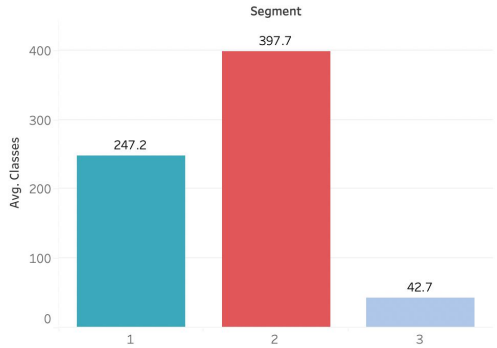



Origination

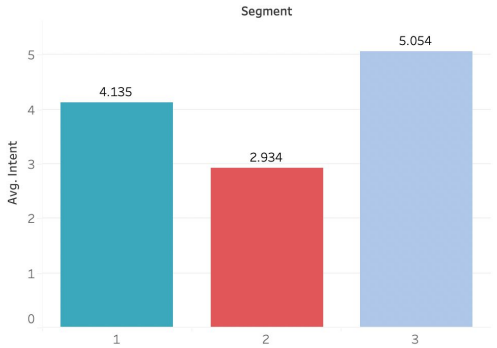



**Overall, segment 2 should be the primary target** because they are the most profitable and have taken the most number of classes. They are the most long-standing customers as well. Their intent and satisfaction are low, thus, they are the most likely to churn. Given the strategic objective of reducing churn rate, retaining this valuable and profitable segment is in Crew's Cup's interest.

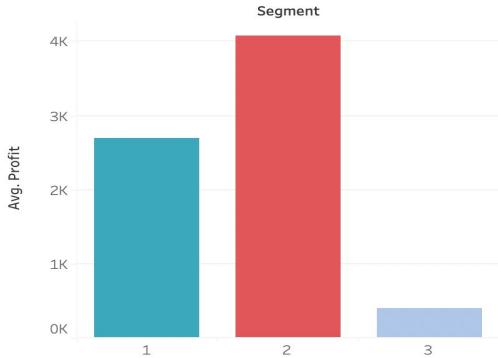
Classes




Intent




profit

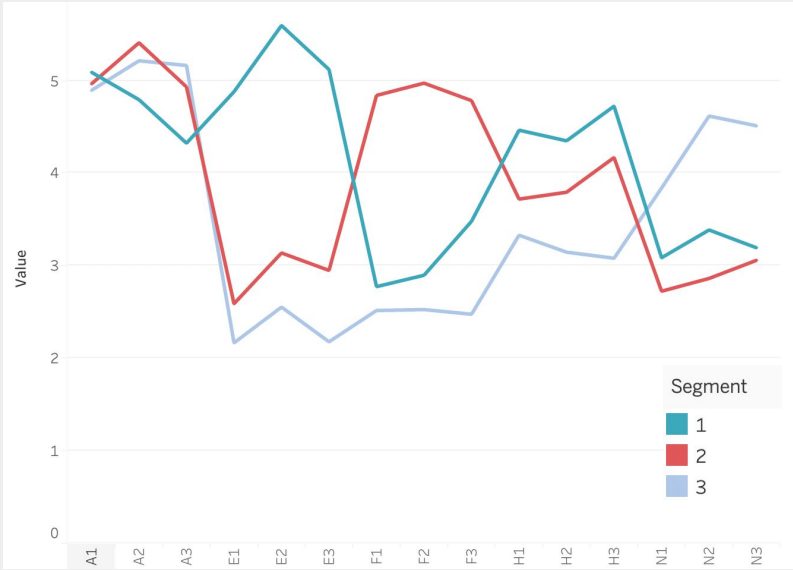


# 07 Segments & Commitment Levels

| Segment | Affective commitment | Normative commitment | Economic commitment | Forced commitment | Habitual commitment |
|---|---|---|---|---|---|
| 1 | 4.73554256 | 3.22157245 | 5.197530864 | 3.050032489 | 4.512020793 |
| 2 | 5.101190476 | 2.881802721 | 2.892006803 | 4.864795918 | 3.893707483 |
| 3 | 5.090196078 | 4.322128852 | 2.299159664 | 2.505882353 | 3.185434174 |

- Segment 1 has a significantly high mean score for economic commitment compared with the other 2 segments and has low normative commitment. This revealed that this segment is not attached to Crew's Cup's values and go there just because of the price and habit, which aligns with the agency's finding that this group is 'less emotionally attached to the brand.' Thus, Monk's campaign may have no effects on them, but the subscription option can attract them and increase retention. However, since they have high satisfaction and intent, and they revealed that they will 'suffer if the relationship was broken' (e2), they are not likely to churn so that's not a problem.
- It is interesting to find that segment 2 has the highest mean of affective commitment score but lowest normative commitment and highest forced commitment, which means they feel pleasure to be a customer, but they are forced to be there and the values of Crew's Cup don't align with them. Since this segment is the most profitable one, this indicates Monk's campaign should target this segment to reduce their likelihood of churn by demonstrating that the brand's values were restored and aligned with this segment's ideals.
- Segment 3 has high affective and normative commitment score. They do not care about the economic benefits and are not forced to go there. The new campaign can reinforce the values they are drawn to and encourage more purchases of this group.

This graph compares commitment components across segments. Segment 1 really cares about price and the will suffer if the relationship was broken (e2). Segment 2 very agrees that Crew's Cup takes care of the customers (a2), but also said they would stop if have other options (f1).



# 08 Regression Model

**Forced commitment has the largest influence on intent, followed by normative and affective, indicating consumers' perceptions towards the brand matters and the new campaign targeting the Faithful should be created.**

Since all the commitment factors are reliable, we can use the average of the factors to calculate the commitment level for each individual in the dataset and build the regression model. Forced committment has the largest influence on intent with a coef of -0.461 and the effect is negative. Normative and affective commitment have a relatively large influence with coef of 0.209 and 0.181. Habitual and economic commitment have the smallest influence, and they are not significant (p-value>0.05). This indicates that whether being forced to choose Crew's Cup affect the intent to purchase the most, and the more forced, the lower the intent to purchase. The economic benefits and habit don't affect consumers' intent to purchase that much and they are bad predictos. Thus, Crew's Cup might not want to create the subscription program for the Budget Hunters and should focus on building brand values and connections with the customers.

| | coef | std err | t | P>\|t\| |
|---|---|---|---|---|
| const | 3.9828 | 0.203 | 19.663 | 0.000 |
| a_avg | 0.1806 | 0.023 | 7.871 | 0.000 |
| n_avg | 0.2091 | 0.023 | 9.005 | 0.000 |
| e_avg | 0.0030 | 0.018 | 0.162 | 0.871 |
| f_avg | -0.4609 | 0.020 | -22.643 | 0.000 |
| h_avg | 0.0190 | 0.024 | 0.809 | 0.419 |

**intent =**
**3.983**
**+ 0.181 * affective**
**+ 0.209 * normative**
**+ 0.003 * economic**
**- 0.461 * forced**
**+ 0.019 * habitual**

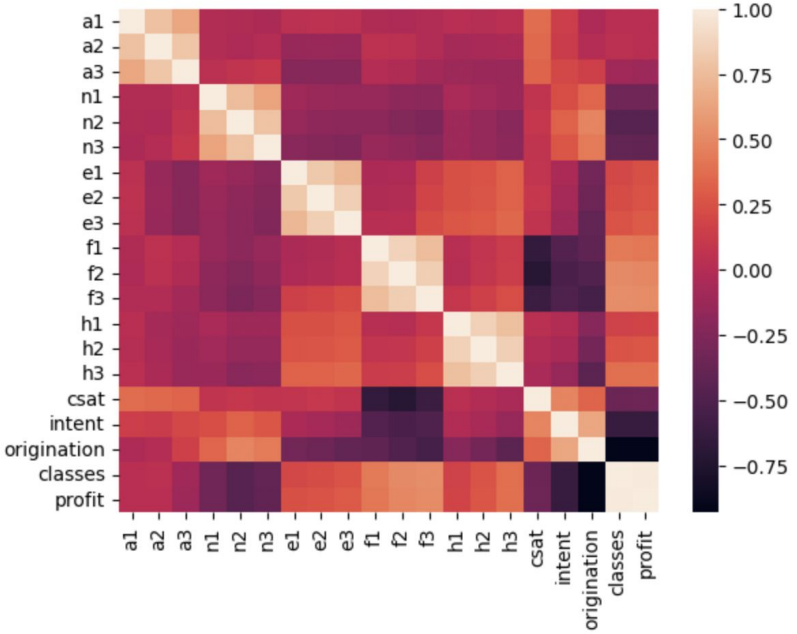
While the regression model helps to identify the most influential factors, there are some **limitations**. The model fit is not really good; there are 2 factors (economic and habitual) that are not significant because of the p-value is > 0.05. Also, the model explains only 34.8% of the data, which means the correlation is not very strong. Thus, while we interpret results from this model, we need to keep in mind that this model is not perfect.

| | |
|---|---|
| R-squared: | 0.348 |
| Adj. R-squared: | 0.346 |
| F-statistic: | 159.8 |
| Prob (F-statistic): | 3.54e-136 |
| Log-Likelihood: | -2211.1 |
| AIC: | 4434. |
| BIC: | 4466. |

# Appendix I

### Correlation matrix of all variables:

```
# Creating a correlation matrix
correlationMatrix  = data.loc[:,numericColumns].corr()
# Printing the correlation matrix.
print(correlationMatrix)
# Displaying the correlation matrix
seaborn.heatmap(correlationMatrix, annot=False)
```




### KMO score:

```
kmo_all,kmo_model=calculate_kmo(data)

kmo_model
```

```
0.7582438886020338
```

The correlation metrix of all variables is plotted to check if there's correlation between variables and if factor analysis can be executed. We can see that there's some significant correlations, so we can find common variables to form to factors.

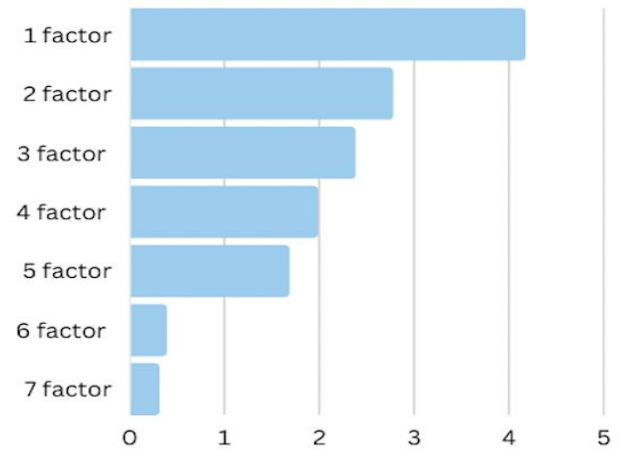
# Appendix II

## Eigenvalue > 1 at 5 factor level:

```
fa = FactorAnalyzer(rotation='varimax')
fa.fit(data)
# Check Eigenvalues
ev, v = fa.get_eigenvalues()
ev
```

here's the cutoff

```
array([4.16054952, 2.76400408, 2.36730912, 1.97250872, 1.67144241,
       0.38064184, 0.30102052, 0.27369491, 0.22333814, 0.2136756 ,
       0.1749329 , 0.14564764, 0.13504264, 0.11577835, 0.1004136 ])
```




## Factor loadings:

```
fa = FactorAnalyzer(n_factors=5,rotation='varimax')
fa.fit(data)
```

```
FactorAnalyzer
FactorAnalyzer(n_factors=5, rotation='varimax', rotation_kwargs={})
```

```
factor_df=pd.DataFrame(fa.loadings_,columns=['Factor1','Factor2','Factor3','Factor4','Factor5'],index=data.columns)
print(factor_df)
```

| | Factor1 | Factor2 | Factor3 | Factor4 | Factor5 |
|---|---|---|---|---|---|
| a1 | -0.025723 | 0.029434 | 0.093649 | 0.814154 | -0.020707 |
| a2 | 0.032380 | -0.026722 | -0.097263 | 0.967840 | -0.024147 |
| a3 | -0.019320 | -0.081647 | -0.182817 | 0.811593 | 0.050016 |
| n1 | -0.092949 | -0.027379 | -0.050663 | 0.003290 | 0.768741 |
| n2 | -0.135120 | -0.084375 | -0.073891 | -0.004392 | 0.960605 |
| n3 | -0.091366 | -0.081801 | -0.158750 | 0.004790 | 0.788065 |
| e1 | -0.007755 | 0.153340 | 0.822711 | -0.063647 | -0.080704 |
| e2 | 0.013018 | 0.133841 | 0.949434 | -0.047035 | -0.102073 |
| e3 | 0.052474 | 0.176444 | 0.842885 | -0.062394 | -0.119315 |
| f1 | 0.879327 | 0.032824 | -0.049558 | 0.009313 | -0.082893 |
| f2 | 0.973922 | 0.025555 | -0.043707 | -0.003860 | -0.115940 |
| f3 | 0.852011 | 0.088202 | 0.158512 | -0.025060 | -0.139967 |
| h1 | -0.012138 | 0.874563 | 0.129150 | -0.028637 | -0.031529 |
| h2 | 0.053297 | 0.954396 | 0.128658 | -0.031384 | -0.068232 |
| h3 | 0.116380 | 0.845300 | 0.215448 | -0.018058 | -0.107603 |

# Appendix III

**Chronbach's Alpha Score of Habitual factor:**

```
datanew.drop(['id','csat','intent','origination','classes','profit','a1','a2','a3','n1','n2','n3','e1','e2','e3','f1','f2','f3'],axis=1,inplace=True)

pg.cronbach_alpha(data=datanew)
```

```
(0.9321936708714786, array([0.926, 0.938]))
```

**Chronbach's Alpha Score of Forced factor:**

```
(0.9319177614554048, array([0.926, 0.938]))
```

**Chronbach's Alpha Score of Economic factor:**

```
(0.921087713693511, array([0.914, 0.928]))
```

**Chronbach's Alpha Score of Affective factor:**

```
(0.895530608858242, array([0.886, 0.904]))
```

**Chronbach's Alpha Score of Normative factor:**

```
(0.886482241743475, array([0.876, 0.896]))
```

# Appendix IV


**Segment differences per segment.** Cell colors indicate to what extent a segment is statistically different from the rest of the population on each segmentation variable.

| | Population | Segment 1 | Segment 2 | Segment 3 |
|---|---|---|---|---|
| **a1** | 4.98 | 5.09 | 4.96 | 4.89 |
| **a2** | 5.12 | 4.79 | 5.41 | 5.21 |
| **a3** | 4.82 | 4.32 | 4.93 | 5.16 |
| **n1** | 3.29 | 3.09 | 2.72 | 3.84 |
| **n2** | 3.74 | 3.39 | 2.86 | 4.62 |
| **n3** | 3.68 | 3.19 | 3.06 | 4.51 |
| **e1** | 3.20 | 4.88 | 2.59 | 2.17 |
| **e2** | 3.75 | 5.59 | 3.14 | 2.55 |
| **e3** | 3.39 | 5.12 | 2.95 | 2.18 |
| **f1** | 3.21 | 2.77 | 4.84 | 2.52 |
| **f2** | 3.29 | 2.90 | 4.97 | 2.53 |
| **f3** | 3.42 | 3.48 | 4.78 | 2.48 |
| **h1** | 3.82 | 4.46 | 3.72 | 3.33 |
| **h2** | 3.73 | 4.35 | 3.79 | 3.15 |
| **h3** | 3.93 | 4.72 | 4.17 | 3.08 |
| **csat** | 3.95 | 4.22 | 3.03 | 4.31 |
| **intent** | 4.19 | 4.13 | 2.93 | 5.05 |
| **origination** | 2 018 | 2 016 | 2 014 | 2 021 |
| **classes** | 205.4 | 247.2 | 397.7 | 42.7 |
| **profit** | 2 143 | 2 699 | 4 073 | 393 |

**Segment description.** Average value of each segmentation variable, overall for each segment (centroid). Segmentation variables that are statistically different from the rest of the population are highlighted in red (lower) or green (higher).