Executive Summary

Objective:

The objective is to increase sales and revenue through the internet. The company decided to utilize the Search Engine Marketing via Google Adsense to increasing the relevance of the company's website. The google Adsense would give company the most "optimal" key words to bid on.

Your Best Model:

The best model used is as follows ;

In this model I have taken the following parameters :

a. Competition
b. Local Monthly Searches
c. Approximate CPC
d. Cost

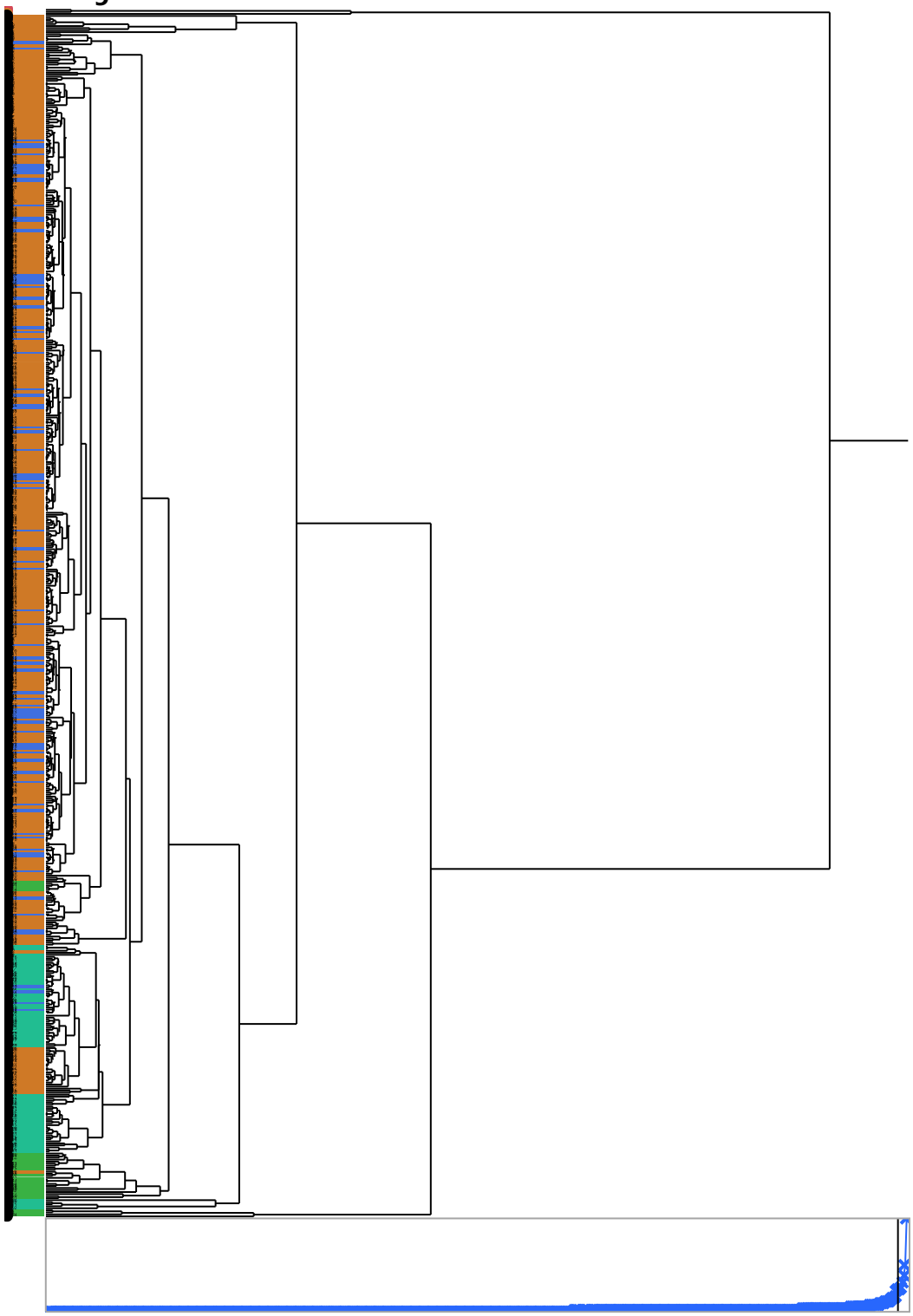
I chose the top three parameters because they were high performing parameters. The rationale behind the final choice of cost is because when plot against CPC would give us insights into estimates number of clicks and also higher the cost and higher the CPC we get a better number of signups and this coincides with adwords with low Cost per Click and low competition but high monthly searches had a high amount of words in common. So we built the model based on this statistical insight.

In this model I have 5 clusters and I omitted the cluster 2 since it was not performing well. For this clustering we went with a hierarchical clustering with ward methodology.

Key Clusters Selected:

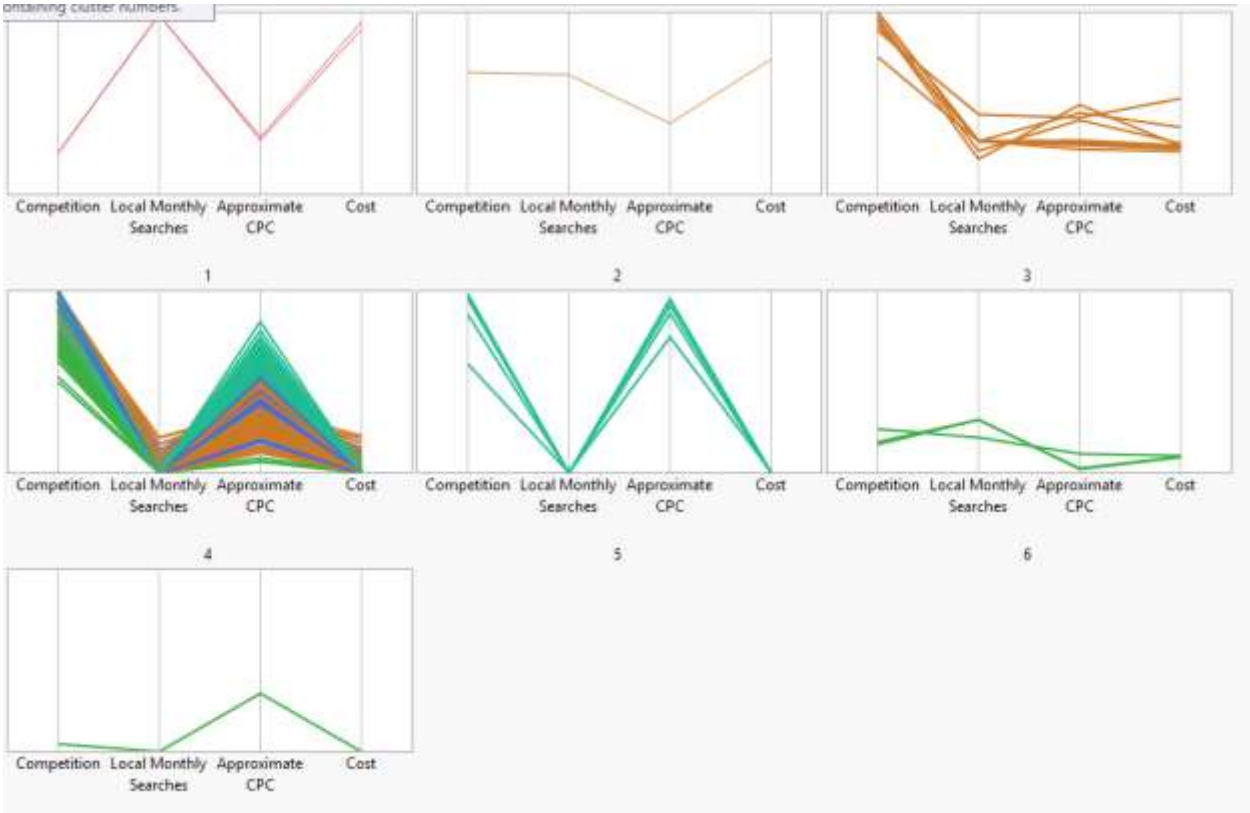
Dendrogram
Hierarchical Clustering
Method = Centroid

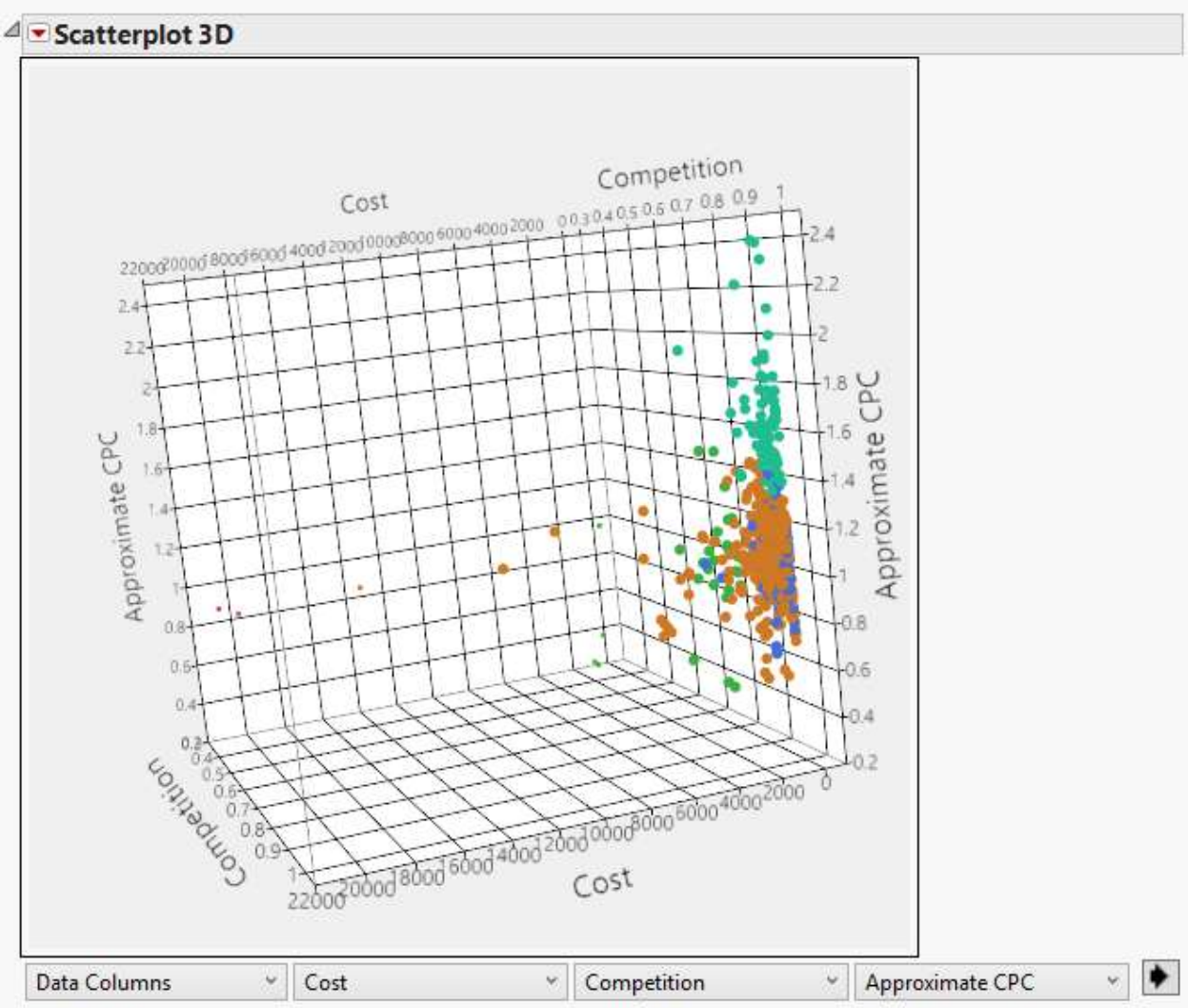
Dendrogram



Key Insights about Clusters:

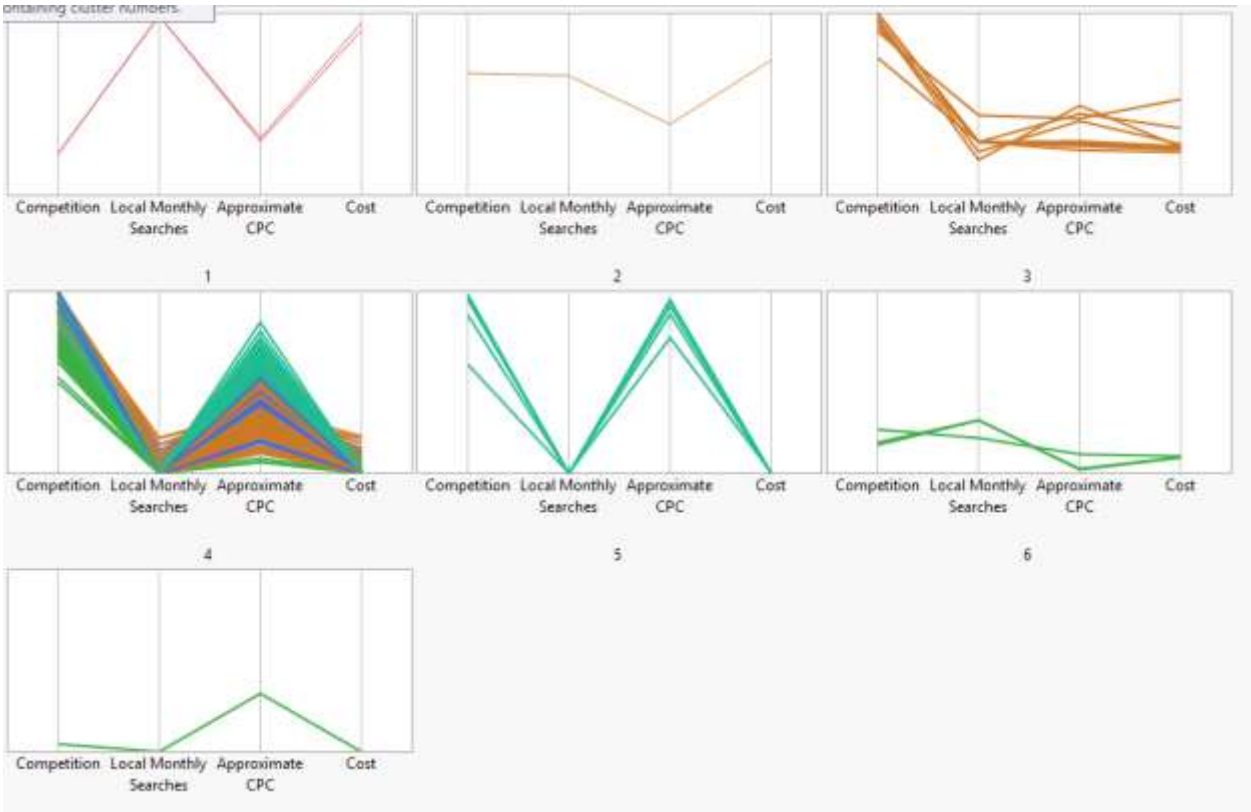
From the below graphs we were able to conclude that our model might have a better model by ignoring 4 , 5 and 7 cluster since it we have least amount of local Monthly searches and the CPC is very high which makes it very undesirable.
The cluster 1 and 2 are the best clusters since they have the best local monthly searches with lesser competition and low CPC. So these clusters have to be included.
The clusters 3 and 6 perform on a average scale in the local monthly searches and workable CPC so we are choosing them.





Why your model is better? (Grill it)

The number of clicks for the model built is approximately 37,000 signups per year. In our model we are trying to find out the one that has high local searches per month and lost CPC since the company would benefit from this. Also we are looking for clusters with low competition so that the bid would be won with a high probability if the company bids on these adwords. When cost is plot against CPC would give us insights into estimates number of clicks and also higher the cost and higher the CPC we get a better number of signups



Conclusion:

With a signups of about 37,000 per year our model is able to ensure that the model would give the company a good insight on how to optimally bid on the google adwords.

Attachments:

Answer the following questions. Provide Answers (don't refer to printouts) and key parts of printouts

1. Build a K-Means Clustering Model to predict the right set of keywords to bid

   MODEL 1

   i) K = 6 (select the number of clusters to be 6)

   ii) Bid on the following Clusters, Cluster1, Cluster5 and Cluster6 – Cluster 1 has large monthly searches, Cluster5 is competitor (Smarty had a party) website (trying to lure competitor customers to your website) and Cluster6 is high competition

   Answer the following questions,

   a) Provide the cluster Means and cluster standard deviations.

Cluster Means

Cluster	Competition	Local Monthly Searches	Approximate CPC
1	0.46	22200	0.94
2	0.9662963	109.777778	1.57342593
3	0.95	6246.66667	1.03266667
4	0.79066667	81.8333333	1.00166667
5	0.4075	4403	0.5175
6	0.96537757	203.826087	1.05908467

Cluster Standard Deviations

Cluster	Competition	Local Monthly Searches	Approximate CPC
1	0	0	0.02
2	0.02983264	236.607517	0.22424425
3	0.06429101	2819.18822	0.15809139
4	0.05390939	93.7194691	0.30207155
5	0.05018715	2689.5403	0.27197197
6	0.02996805	451.298175	0.14793648

b) Interpret Cluster1, Cluster5 and Cluster 6 based on the means

Cluster 1 is broad and has high mean of monthly searches and low average competition.
Cluster 5 focuses on taking business away from a competitor.
Cluster 6 is highly competitive with low searches and high cost per connect.

c) What is the net profit

-$3,950

d) What is the estimated signups per month

2,874

e) What is the estimated signups per year

34,482

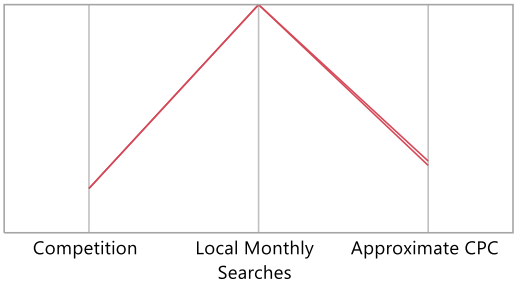
f) How many estimated monthly clicks are allowed
21,407

g) How many estimated monthly clicks are allowed if there is no budget constraints
22,663

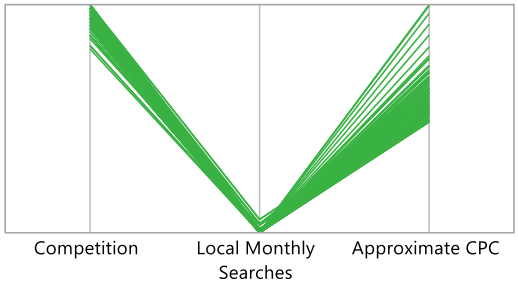
	#Impressions	Avg. CPC
Available Search Volume	259095	$1.00
Available Head KV	127300	$0.93
Available Mid&Tail KV	131795	$1.07
Selected KV	**151084**	**$0.93**
Est. Impressions Share(%)	75%	
Est. Impressions	113,313	

Monthly Budget	$20,000
Revenue per Conversion	$15
Est. CTR%	20.00%
Monthly Clicks if no budget constraint	22,663
Monthly Clicks allowed	21407
Actual CPC	$0.93
Assumed CR	5.00%
CPC Scale	15%
CR Scale (Non Linear)	10%
Amount Spent	$19,999
Net Profit	**($3,950)**
Estimated Signups per Month	**2,874**
Estimated Signups per Year	**34,482**

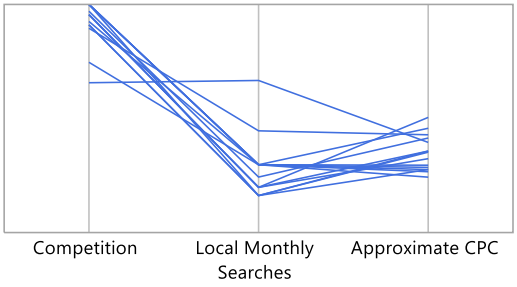
Parallel Coordinate Plot



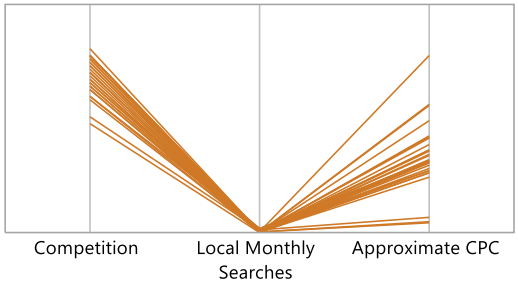
1



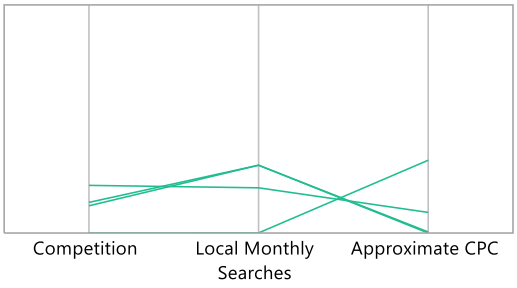
2



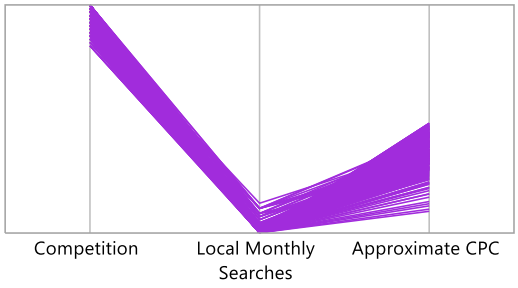
3



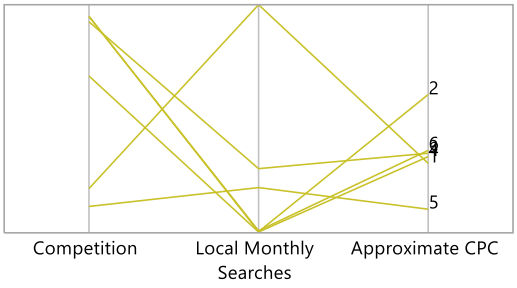
4



5



6



Cluster Means

2. Rebuild a K-Means Clustering Model to predict the right set of keywords to bid

   MODEL 2

   i) K = 5 (select the number of clusters to be 5)

   ii) Bid on the following Clusters, Cluster1, Cluster2 and Cluster3

Answer the following questions,

a) Provide the cluster Means and cluster standard deviations

Cluster Means

Cluster	Competition	Local Monthly Searches	Approximate CPC
1	0.56333333	19733.3333	1.00333333
2	0.9568709	199.275711	1.05035011
3	0.96285714	5635.71429	1.02571429
4	0.95982759	103.913793	1.56362069
5	0.49166667	2978.66667	0.58

Cluster Standard Deviations

Cluster	Competition	Local Monthly Searches	Approximate CPC
1	0.1461354	3488.39345	0.09104334
2	0.047295	442.289204	0.15189787
3	0.04414725	1707.80023	0.16140961
4	0.04605062	229.408368	0.22558505
5	0.12601808	2980.85595	0.28378396

b) Interpret Cluster1, Cluster2 and Cluster 3 based on the means
Cluster 1 is considered broad searches, but now are more competitive than the other 2 keywords since it includes "party cups"
Cluster 2 is high competition, low monthly searches, and medium low CPC, so the searches are more related to tablewear but not necessarily for specific parties.
Cluster 3 exhibits high competition, medium low monthly searches, and medium CPC, making these so the searches are more related to party.

c) What is the net profit
-$5,345

d) What is the estimated signups per month
2,557

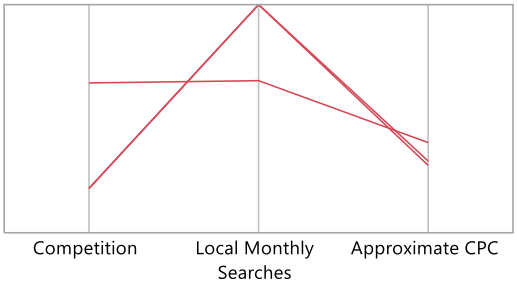
e) What is the estimated signups per year
30,680

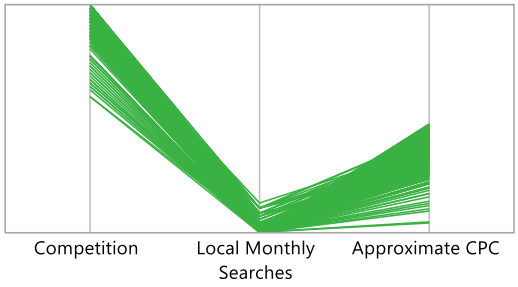
f) How many estimated monthly clicks are allowed
19,547

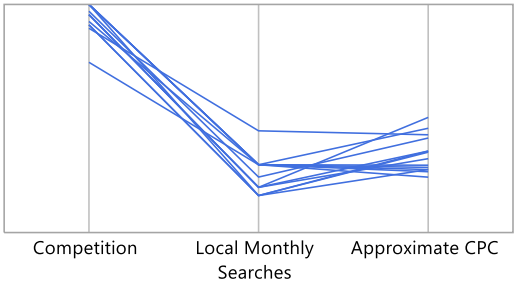
g) How many estimated monthly clicks are allowed if there is no budget constraints
34,375

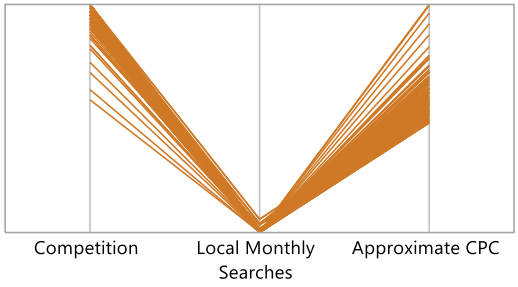
	#Impressions	Avg. CPC
Available Search Volume	259095	$1.00
Available Head KV	127300	$0.93
Available Mid&Tail KV	131795	$1.07
Selected KV	**229169**	**$1.02**
Est. Impressions Share(%)	75%	
Est. Impressions	171,877	
Monthly Budget	$20,000	
Revenue per Conversion	$15	
Est. CTR%	20.00%	
Monthly Clicks if no budget constraint	34,375	
Monthly Clicks allowed	19547	
Actual CPC	$1.02	
Assumed CR	5.00%	
CPC Scale	15%	
CR Scale (Non Linear)	10%	
Amount Spent	$19,999	
Net Profit	**($5,345)**	
Estimated Signups per Month	**2,557**	
Estimated Signups per Year	**30,680**	

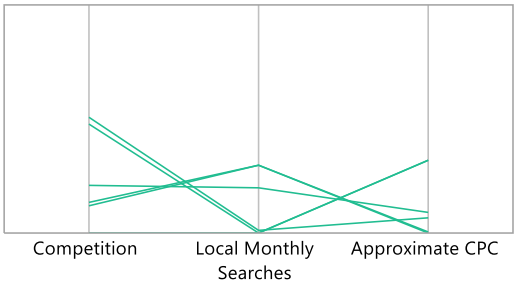
Parallel Coordinate Plot

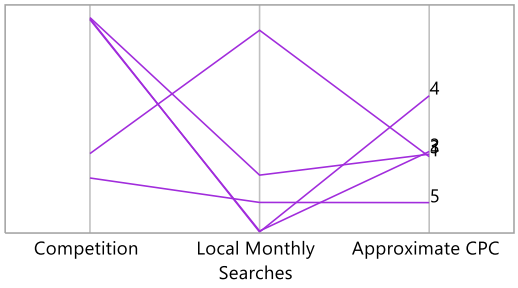












3. Build a Hierarchical Clustering Model to predict the right set of keywords to bid

MODEL 3

i) K = 6 (select the number of clusters to be 6)

ii) Bid on the following Clusters, Cluster1, Cluster3 and Cluster4 – Cluster 1 has large monthly searches etc.,

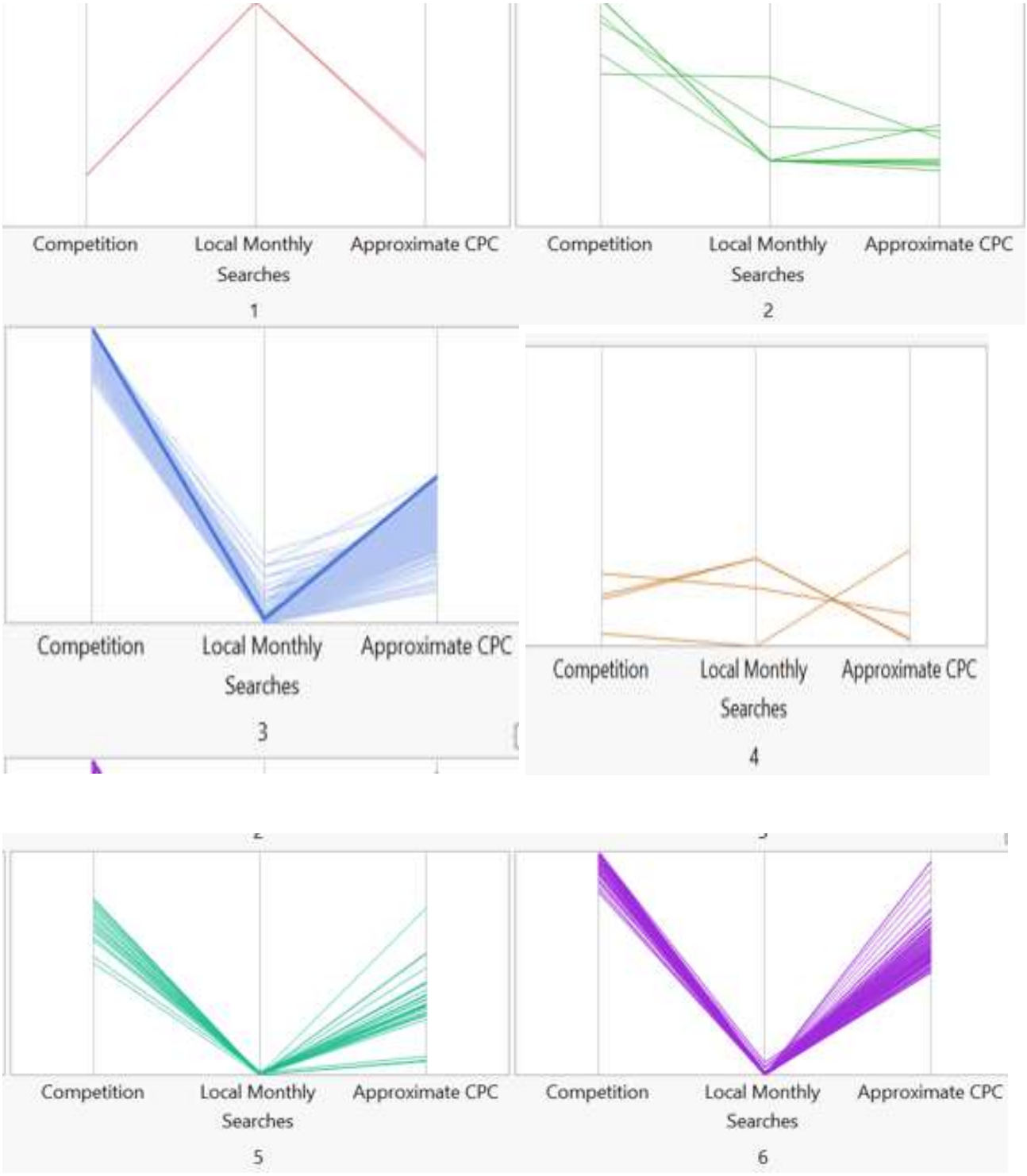
Answer the following questions,

a) What is the net profit
   $4,430
b) What is the estimated signups per month
   2,701
c) What is the estimated signups per year
   32,414
d) How many estimated monthly clicks are allowed
   20,765
e) How many estimated monthly clicks are allowed if there is no budget constraints

27,371

	#Impressions	Avg. CPC
Available Search Volume	259095	$1.00
Available Head KV	127300	$0.93
Available Mid&Tail KV	131795	$1.07
Selected KV	**182471**	**$0.96**
Est. Impressions Share(%)	75%	
Est. Impressions	136,853	
Monthly Budget	$20,000	
Revenue per Conversion	$15	
Est. CTR%	20.00%	
Monthly Clicks if no budget constraint	27,371	
Monthly Clicks allowed	20765	
Actual CPC	$0.96	
Assumed CR	5.00%	
CPC Scale	15%	
CR Scale (Non Linear)	10%	
Amount Spent	$19,999	
Net Profit	**($4,430)**	
Estimated Signups per Month	**2,701**	
Estimated Signups per Year	**32,414**	

Parallel Coordinate Plot



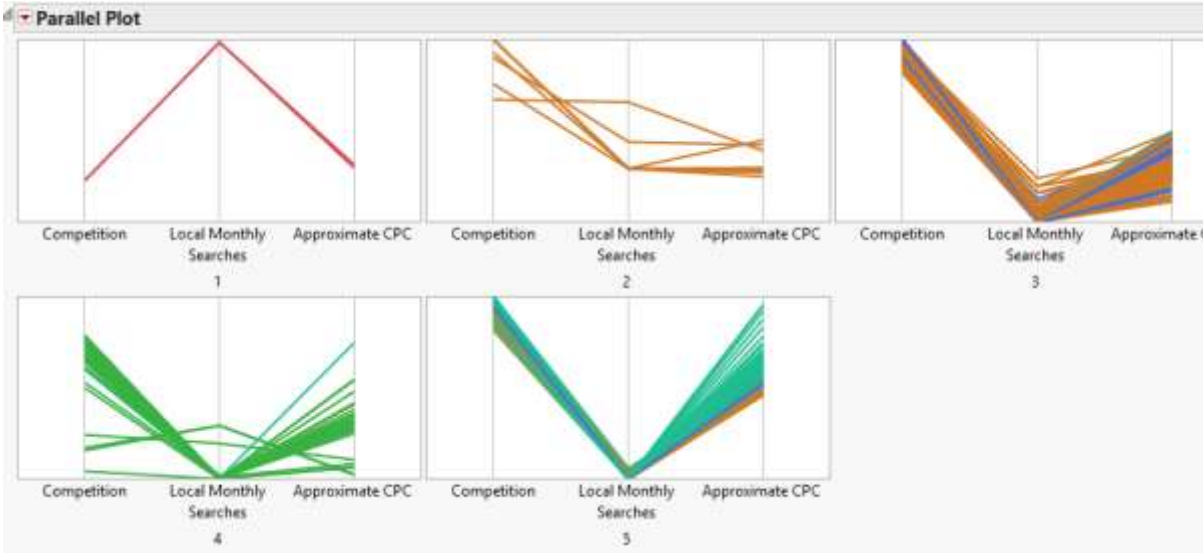
4. Adjust the Hierarchical Clustering Model to predict the right set of keywords to bid
   MODEL 4
   i) K = 5 (select the number of clusters to be 5)
   ii) Bid on the following Clusters, Cluster1, Cluster2 and Cluster3

Answer the following questions,

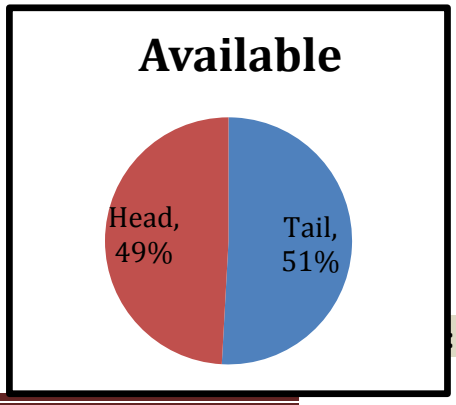
a) What is the net profit
b) What is the estimated signups per month
c) What is the estimated signups per year
d) How many estimated monthly clicks are allowed
e) How many estimated monthly clicks are allowed if there is no budget constraints

Answer the following questions,

a) What is the net profit
   $5390
b) What is the estimated signups per month
   2557
c) What is the estimated signups per year
   30679
d) How many estimated monthly clicks are allowed
   19477
e) How many estimated monthly clicks are allowed if there is no budget constraints
   34374



	#Impressions	Avg. CPC
Available Search Volume	259095	$1.00
Available Head KV	127300	$0.93
Available Mid&Tail KV	131795	$1.07
Selected KV	**229159**	**$1.03**
Est. Impressions Share(%)	**75%**	



Est. Impressions	**171,869**	

Scroll down for legends explanation

Break-even Analysis (First Month)

Monthly Budget	$20,000
Revenue per Conversion	$15
Est. CTR%	20.00%
Monthly Clicks if no budget constraint	34,374
Monthly Clicks allowed	19477
Actual CPC	$1.03
Assumed CR	5.00%
CPC Scale	15%
CR Scale (Non Linear)	10%

CR%	Amount Spent
0.5%	$20,000
1.0%	$20,000
2.5%	$20,000
5.0%	$20,000
10.0%	$20,000
15.0%	$20,000
20.0%	$20,000
25.0%	$20,000

Amount Spent	$20,000
Net Profit	**($5,390)**
Estimated Signups per Month	**2,557**
Estimated Signups per Year	**30,679**

5. Build your own best Clustering Model, to maximize the Number of Signups.
   a) Explain your Model
   b) Expected Profit
   c) Expected Number of Signups

Answers :

a) In this model I have taken the following parameters :
   a. Compeition
   b. Local Monthly Searches

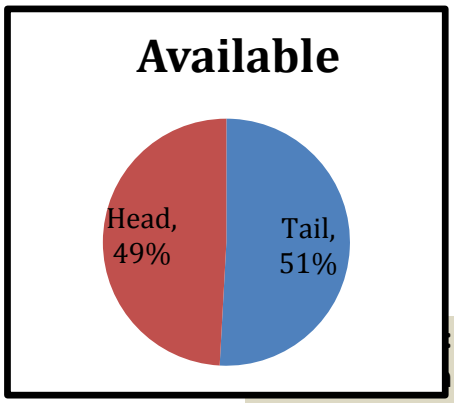
c. Approximate CPC
d. Cost

I chose the top three parameters because they were high performing parameters. The rationale behind the final choice of cost is because when plot against CPC would give us insights into estimates number of clicks and also higher the cost and higher the CPC we get a better number of signups and this coincides with adwords with low Cost per Click and low competition but high monthly searches had a high amount of words in common. So we built the model based on this statistical insight.

In this model I have 5 clusters and I omitted the cluster 2 since it was not performing well.

b) Estimated Profit : $4685
c) Expected number of signups: 36984

	#Impressions	Avg. CPC
Available Search Volume	259095	$1.00
Available Head KV	127300	$0.93
Available Mid&Tail KV	131795	$1.07
Selected KV	**140619**	**$0.93**
Est. Impressions Share(%)	**75%**	
Est. Impressions	**105,464**	



for legends explanation

Break-even Analysis (First Month)

Monthly Budget	$20,000
Revenue per Conversion	$15
Est. CTR%	20.00%
Monthly Clicks if no budget constraint	21,093
Monthly Clicks allowed	21493
Actual CPC	$0.93
Assumed CR	5.00%

CR%	Amount Spent
0.5%	$19,627
1.0%	$19,627
2.5%	$19,627
5.0%	$19,627
10.0%	$19,627
15.0%	$19,627

CPC Scale	15%
CR Scale (Non Linear)	10%

20.0%	$19,627
25.0%	$19,627

Amount Spent	$20,000
Net Profit	**($4,175)**
Estimated Signups per Month	**3,079**
Estimated Signups per Year	**36,948**