Data Analytics Lab 3

---Exercise 1---

Contingency Matrixes/tables (k=64)

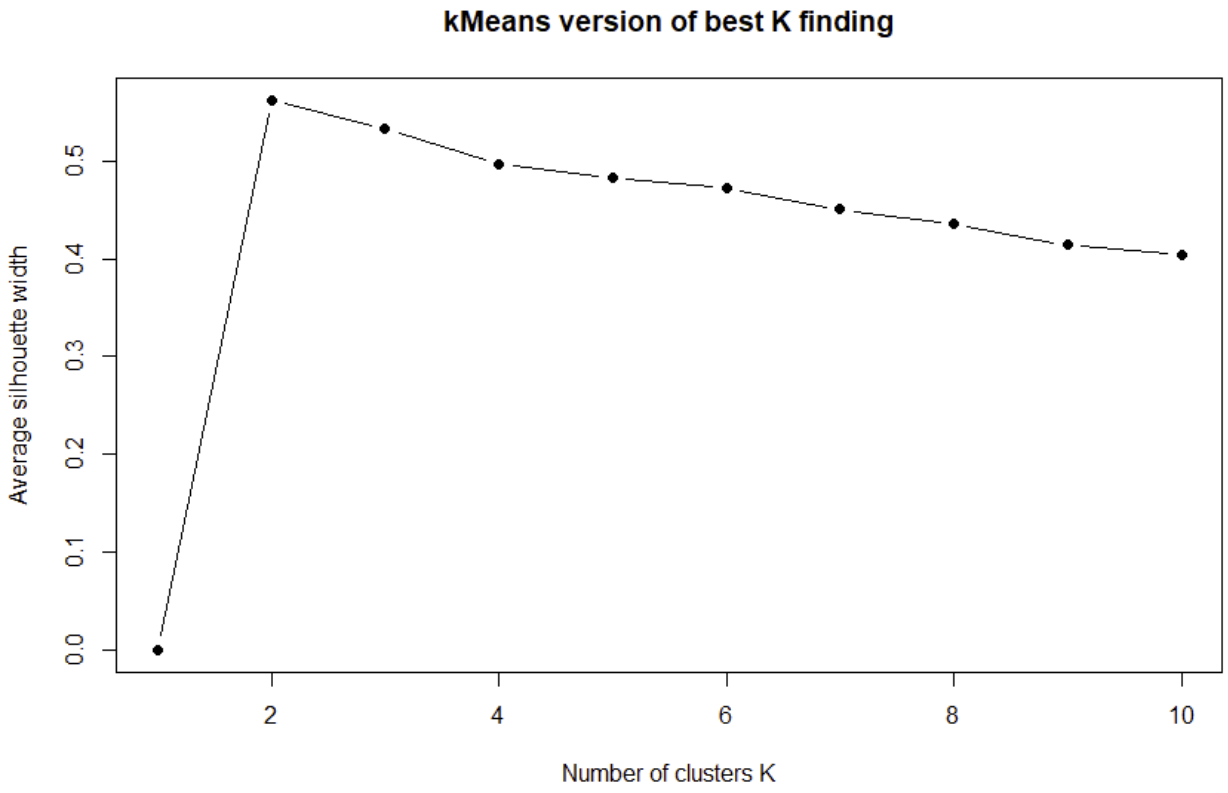
```
> ##Weight has better accuracy.
> table(knn.predicted_size, abalone.test[,10], dnn=list('size_knn','actual'))
        actual
size_knn adult old young
   adult   376 190    98
   old      36  55     3
   young    88  36   294
> table(knn.predicted_weight, abalone.test[,10], dnn=list('weight_knn','actual'))
          actual
weight_knn adult old young
     adult   393 166    93
     old      33  99     1
     young    74  16   301
```

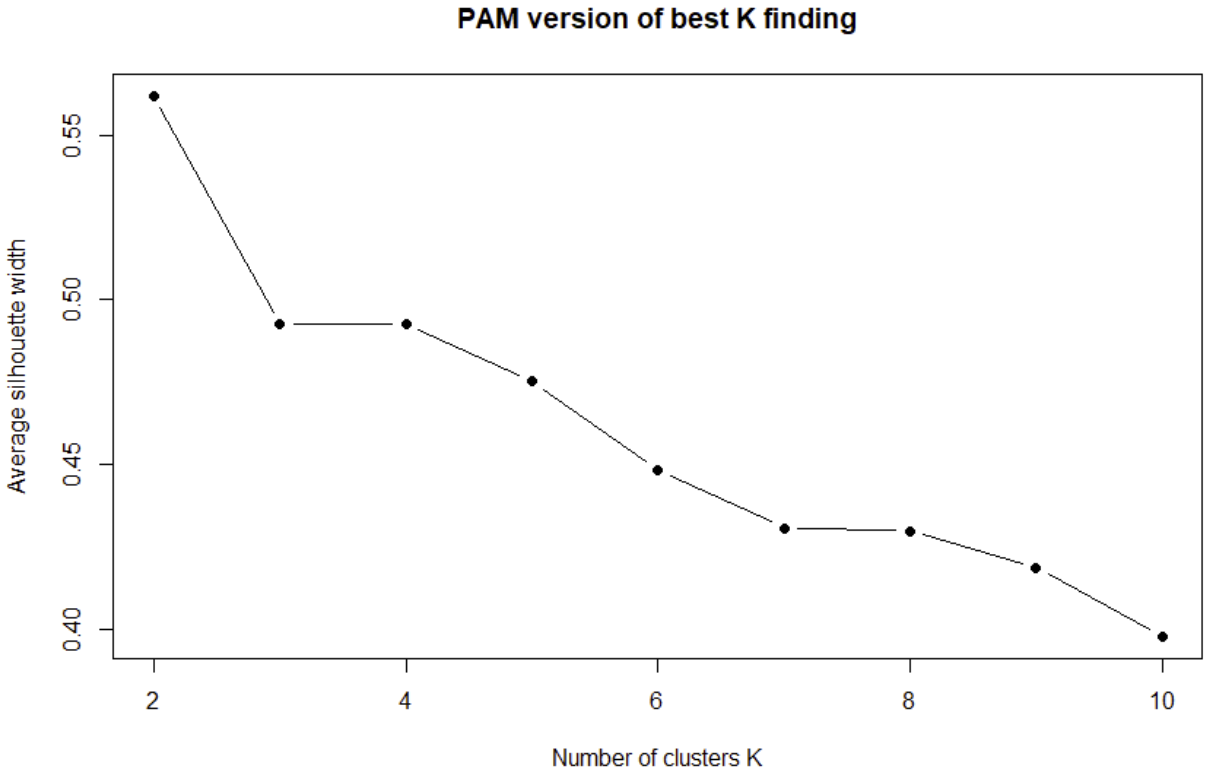
Best k for knn in the range (32-128)

```
> print(max_k_accuracy)
 [1] 0.6845238 0.6921769 0.6930272 0.6930272 0.6921769 0.6964286 0.6947279 0.6938776 0.6972789 0.6930272 0.6913265 0.6921769 0.6887755 0.6870748
[15] 0.6887755 0.6930272 0.6930272 0.6921769 0.6921769 0.6904762 0.6904762 0.6862245 0.6947279 0.6870748 0.6913265 0.6887755 0.6887755 0.6904762
[29] 0.6879252 0.6887755 0.6845238 0.6819728 0.6845238 0.6828231 0.6862245 0.6819728 0.6862245 0.6836735 0.6853741 0.6845238 0.6819728 0.6794218
[43] 0.6802721 0.6836735 0.6785714 0.6751701 0.6802721 0.6760204 0.6785714 0.6802721 0.6794218 0.6760204 0.6768707 0.6768707 0.6794218 0.6768707
[57] 0.6785714 0.6751701 0.6768707 0.6768707 0.6777211 0.6768707 0.6760204 0.6777211 0.6768707 0.6709184 0.6743197 0.6768707 0.6777211 0.6760204
[71] 0.6777211 0.6777211 0.6743197 0.6768707 0.6768707 0.6768707 0.6751701 0.6709184 0.6726190 0.6709184 0.6700680 0.6683673 0.6700680 0.6700680
[85] 0.6700680 0.6700680 0.6700680 0.6726190 0.6709184 0.6734694 0.6726190 0.6683673 0.6666667 0.6683673 0.6692177 0.6700680 0.6734694
> print("The best k for knn: ")
[1] "The best k for knn: "
> print(which.max(max_k_accuracy) + 31)
[1] 40
```

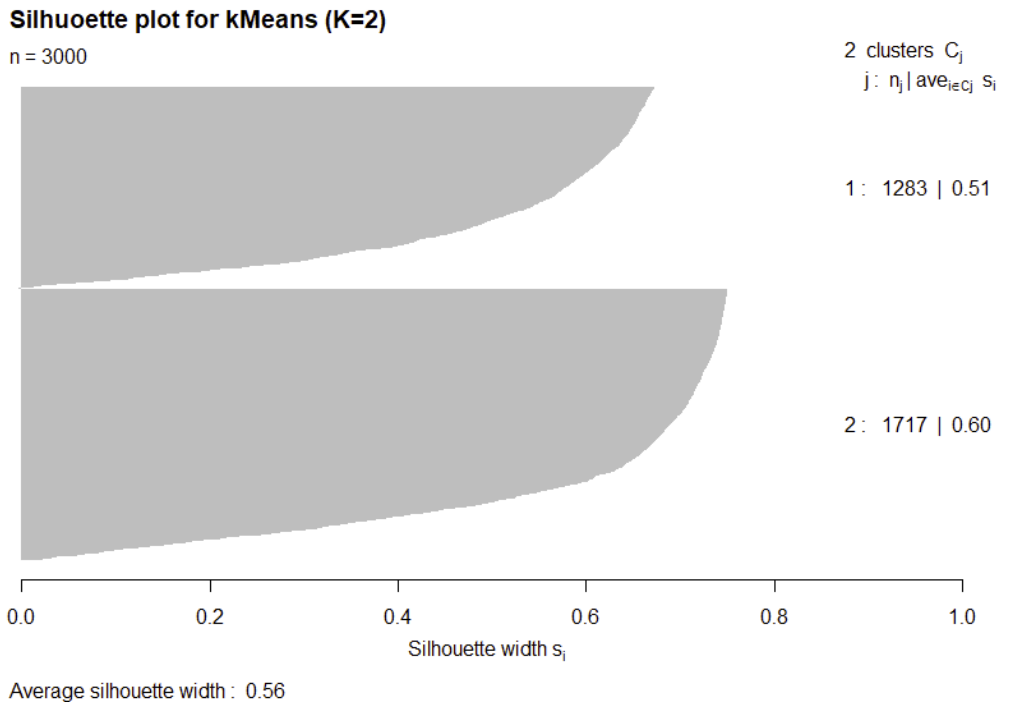
Best k found is 40.

---Exercise 2---

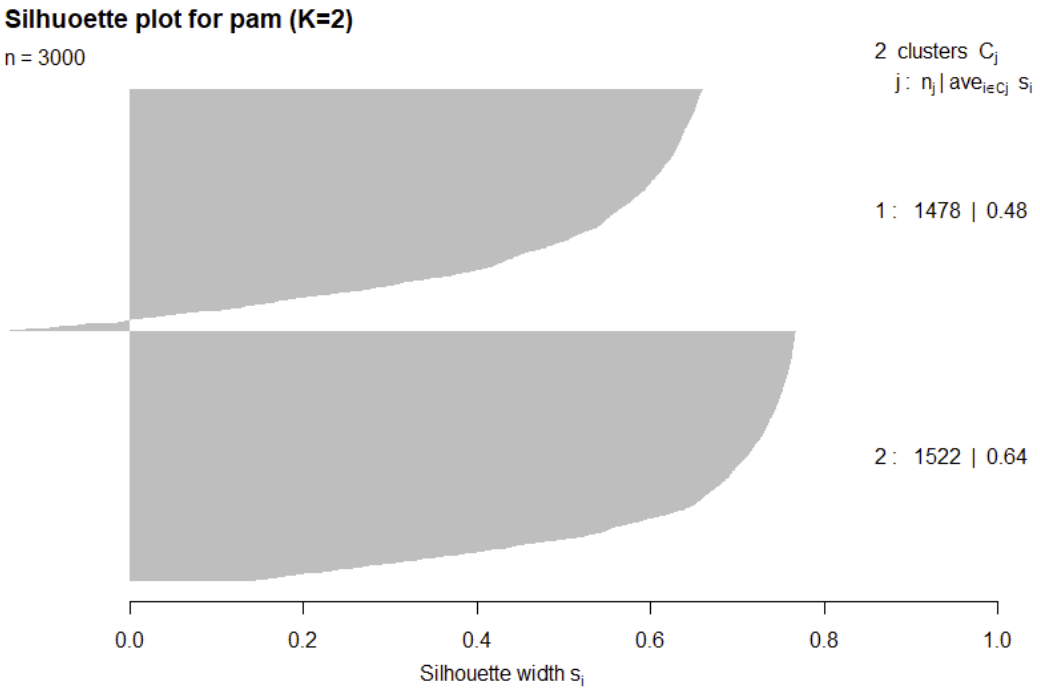

kMeans version of best K finding
Average silhouette width
0.0
0.1
0.2
0.3
0.4
0.5
2
4
6
8
10
Number of clusters K


PAM version of best K finding
Average silhouette width
0.40
0.45
0.50
0.55
2
4
6
8
10
Number of clusters K

Silhuotte plots:



Average silhouette width : 0.56



Average silhouette width : 0.56