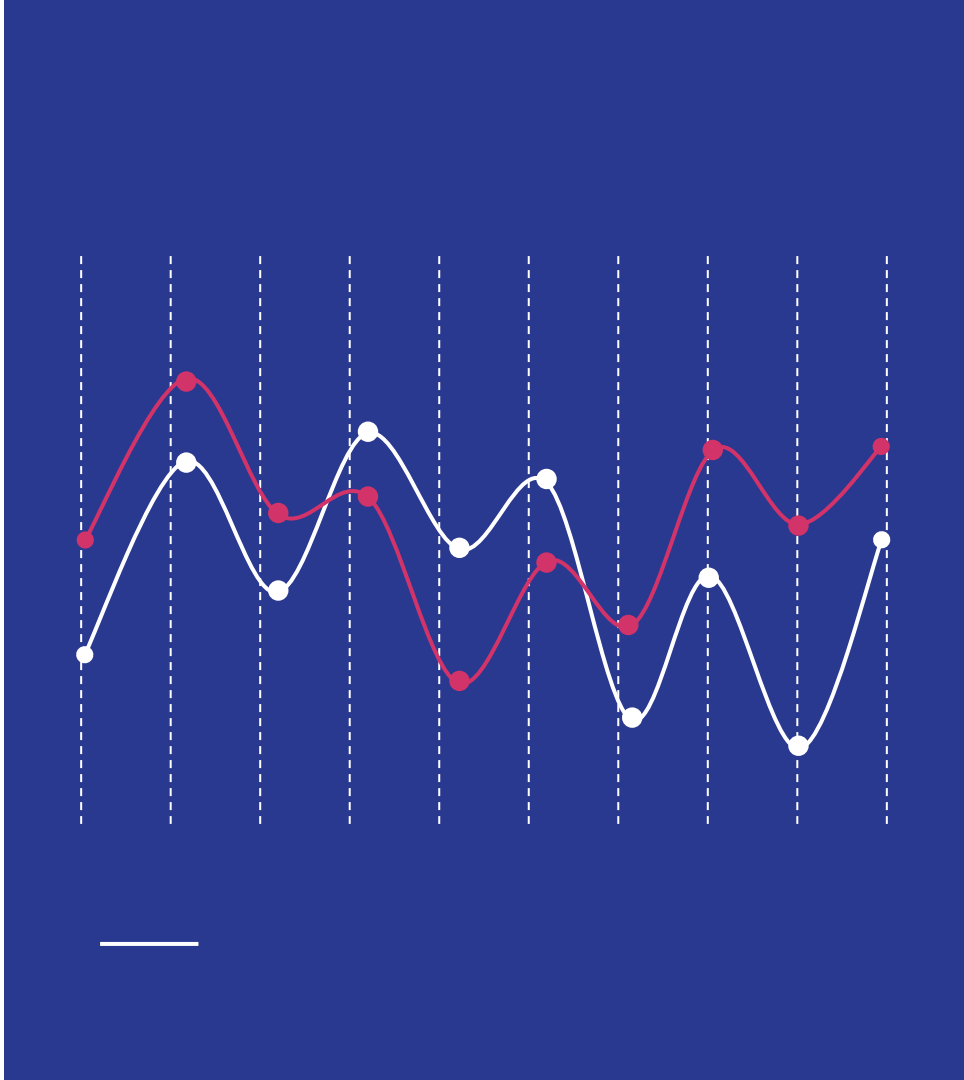
Podcast listeners client base demographic segmentation using advanced Analytics

By Omar Hassan



What is the research Question

What are the various podcast listeners characteristic, and demographics segments in the podcast space?

Challenges

Multivariable data

Our Data has over 500 columns. Our solution was to do some cleaning. Dropped columns, fill empty fields and segregate data into numeric and categorical data frames

Categorical data

Our data contains a lot of categorical data. It is really challenging to get such data to talk to models that reads only numbers. So, we used encoding method to convert the data to ones and zeros

Interpretation

Interpreting the results of machine learning advanced analytical models is not an easy task. However, we leveraged some visualization techniques and joining of data points to produce insights

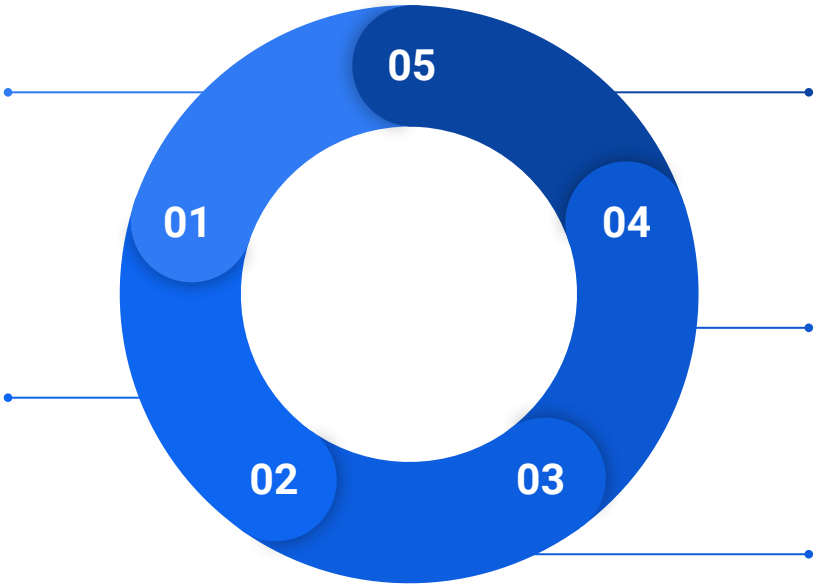
The Process

Data Cleaning & encoding

At this step we cleaned the data, and converted our categorical data into 1 and 0 so models can read it.

Data scaling

At this step we scaled our new data so the mean is 0 and standard deviation is 1 for all our variables



Interpretation & insights

Here we leveraged client unique ID within each cluster and mapped it back to our original data so we can drive insight on each cluster

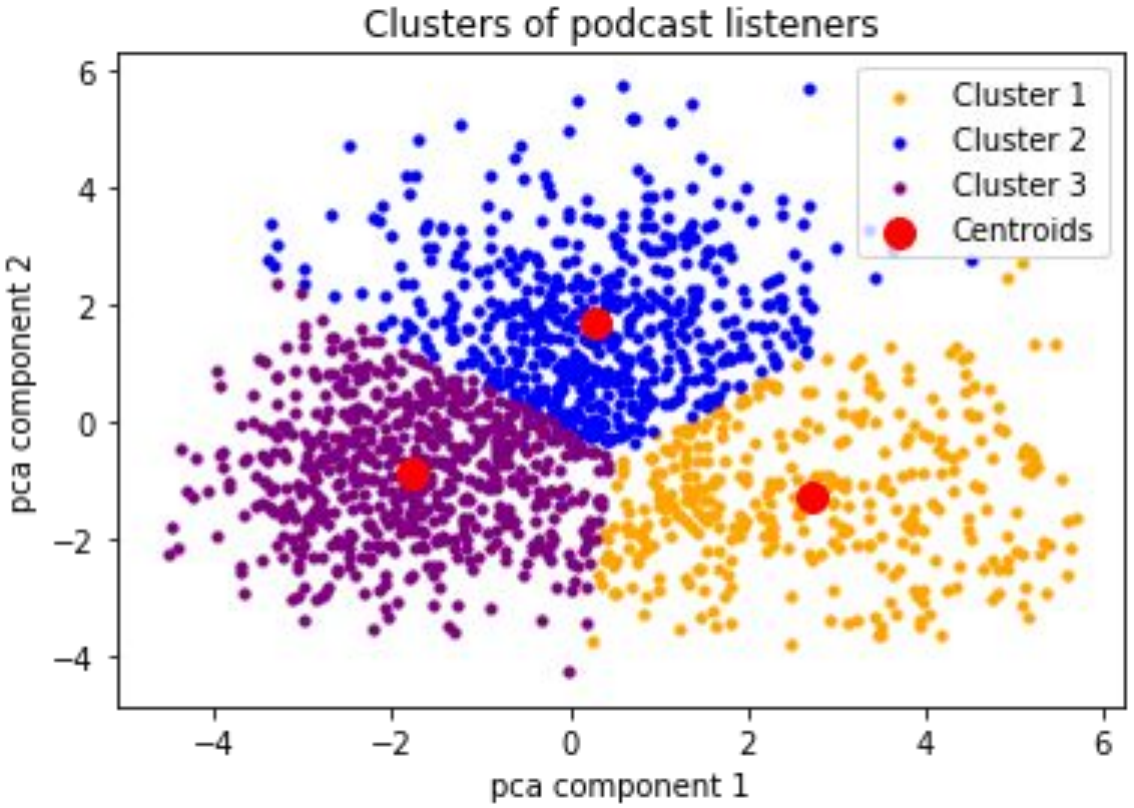
Kmeans clustering

Here we determined the optimal number of clusters and passed that along with our new reduced data set into Kmeans algorithm to create 3 clusters of data points

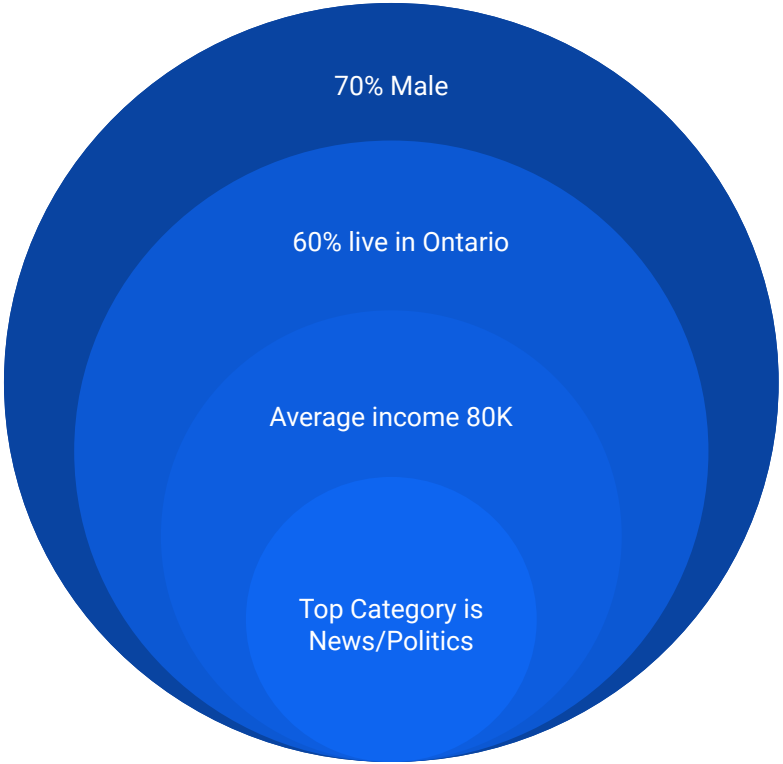
Feature Reduction

Here we used unsupervised algorithm called Principal component analysis to reduce our multivariable data set into just 2 columns while retaining most of the information.

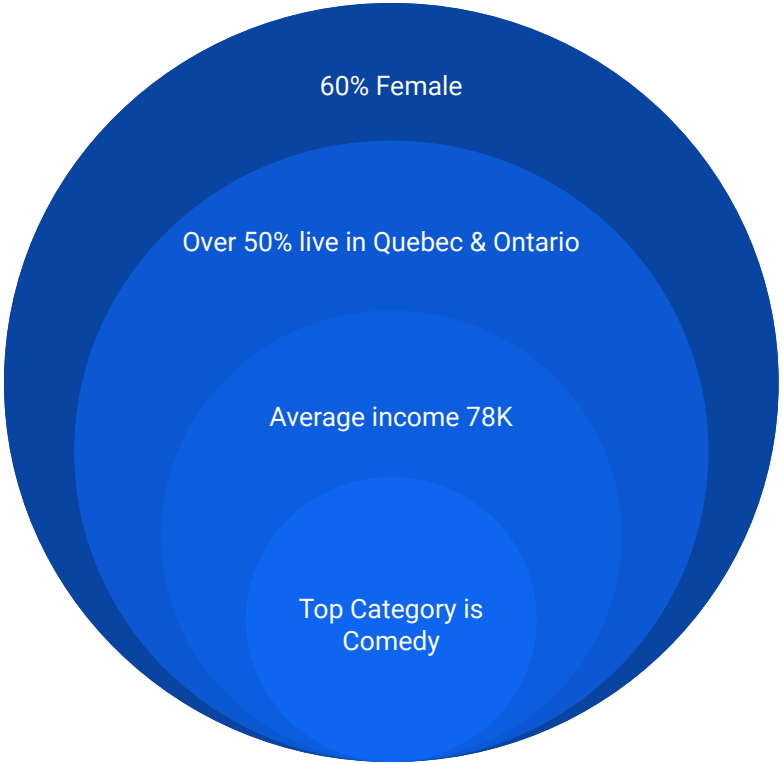
Clusters


Clusters of podcast listeners
Cluster 1
Cluster 2
Cluster 3
Centroids
pca component 2
pca component 1
6
4
2
0
−2
−4
−4
−2
0
2
4
6

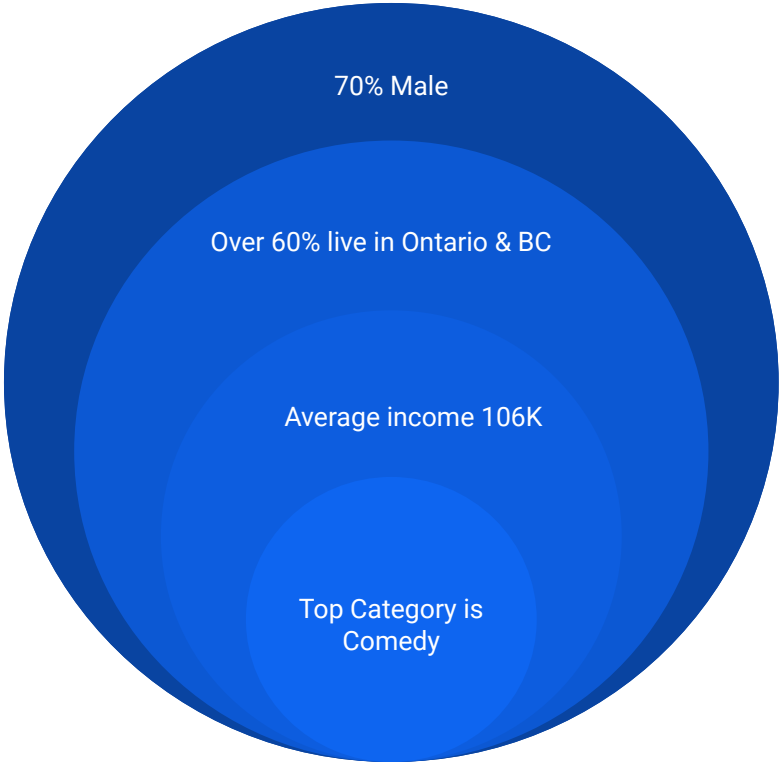
The results - Cluster 1



The results - Cluster 2



The results - Cluster 3



Conclusion

01	**High income Segment**	• High Income • Male • Young in 30s • Listen to Comedy
02	**Female/Quebec Segment**	• Average income • Female • Quebec • Listen to Comedy
03	**Older age segment**	• Male • Average income • Above 60 years old • Listen to Politics & news