

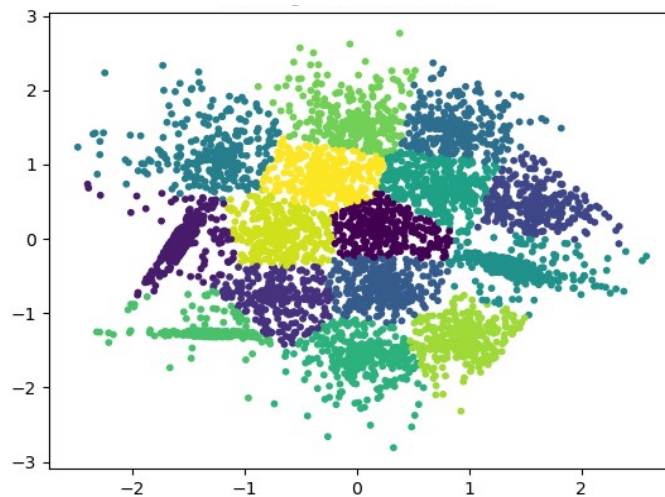
# UNIVERSIDAD POLITÉCNICA DE VICTORIA

## INTELIGENCIA DE NEGOCIOS

DR. MARIO GARZA FABRE

### PROYECTO UNIDAD 4

#### CLUSTERING



#### INTEGRANTES:

ARMANDO ISAAC HERNÁNDEZ MUÑIZ

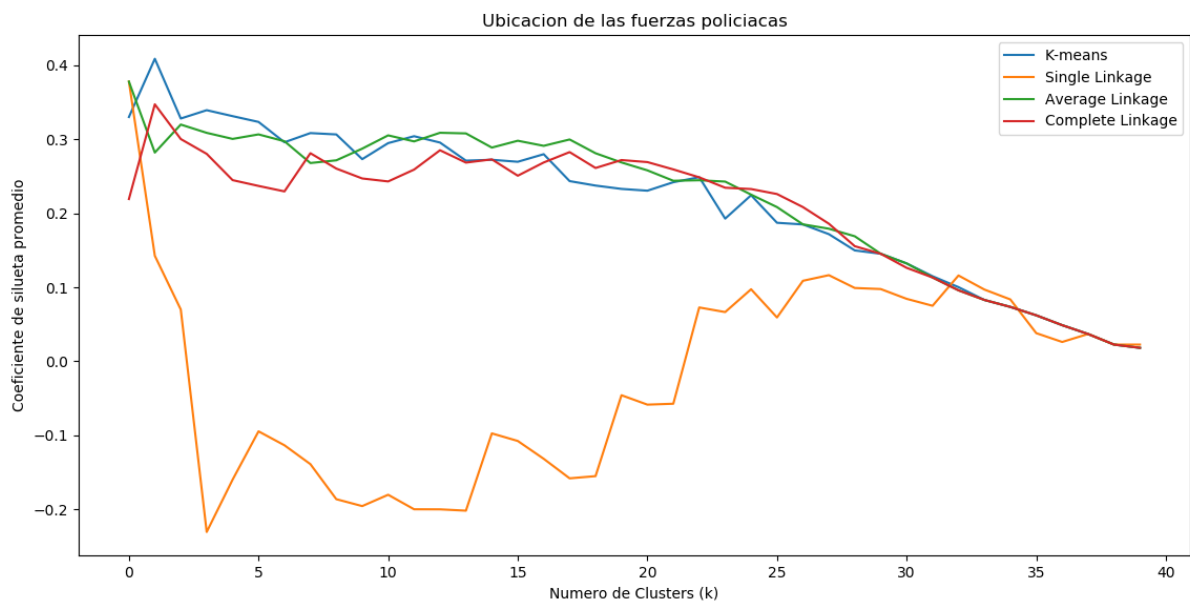
NORBERTO III MORALES AVITIA

GABRIELA ROBLES RODRÍGUEZ

**JULIO 2018**

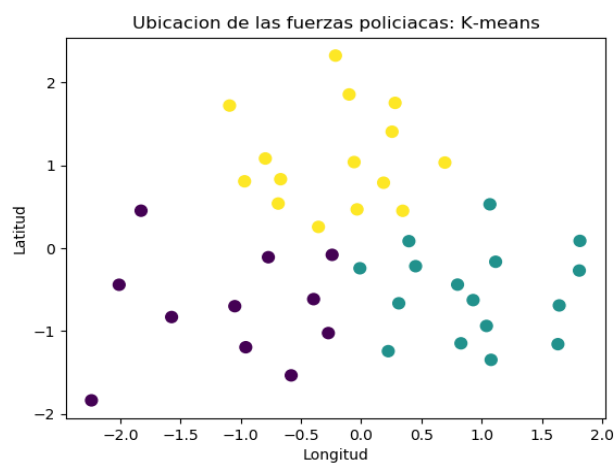
Visualización de resultados.

Se presentaran los resultados del conjunto de datos de “Ubicación de las fuerzas policiacas”

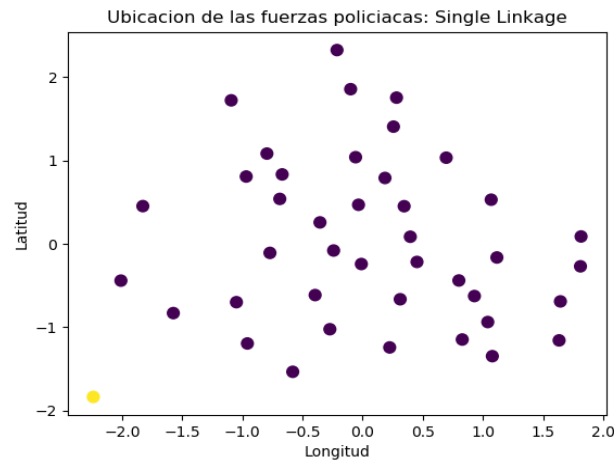


En esta grafica se puede observar el valor de coeficiente de silueta aplicada a cada algoritmo de clustering para diferentes valores de clusters (K), y vemos que en la aplicación de “Single Linkage” el valor del coeficiente decrementa mucho cuando K esta entre el los valores [2, 5] y despues vuelve e subir considerablemente superando en algun momento a los otros tres algoritmos, los cuales obtienen coeficientes muy similares para cada uno de los valores de K.

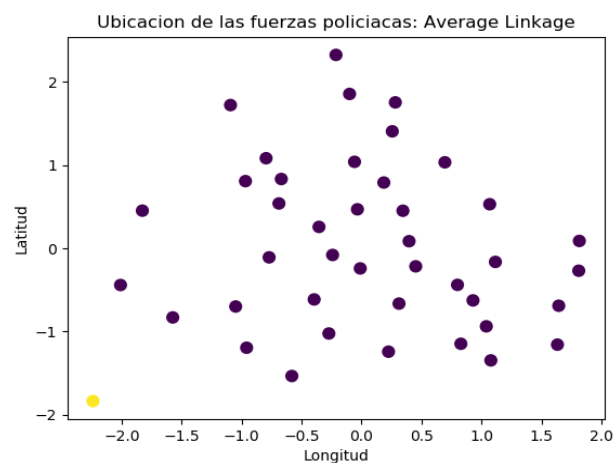
Graficas de dispersión generadas con los valores de K de acuerdo al mejor coeficiente de cada algoritmo. (Conjunto de datos con 2 atributos)



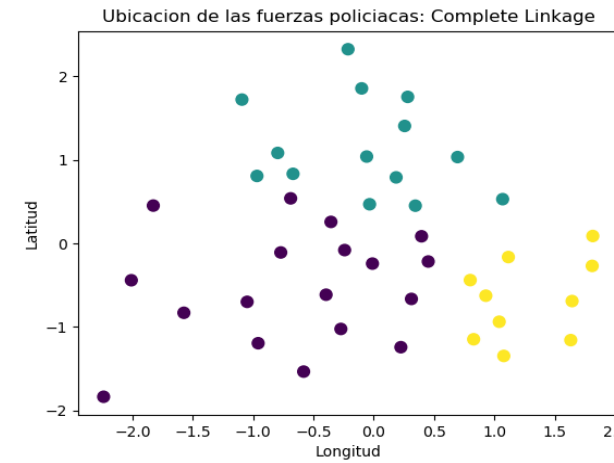
K=3 Silhouette\_score=0.408



K=2 Silhouette\_score=0.377

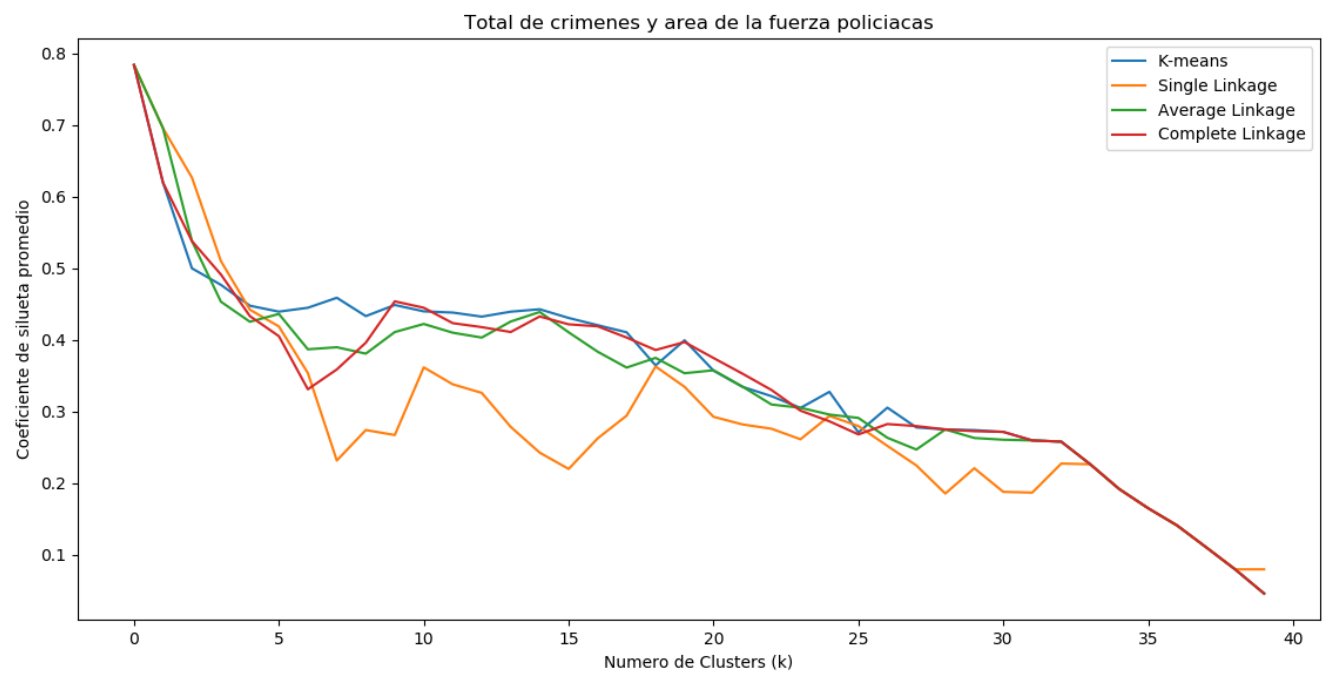


K=2 Silhouette\_score=0.377



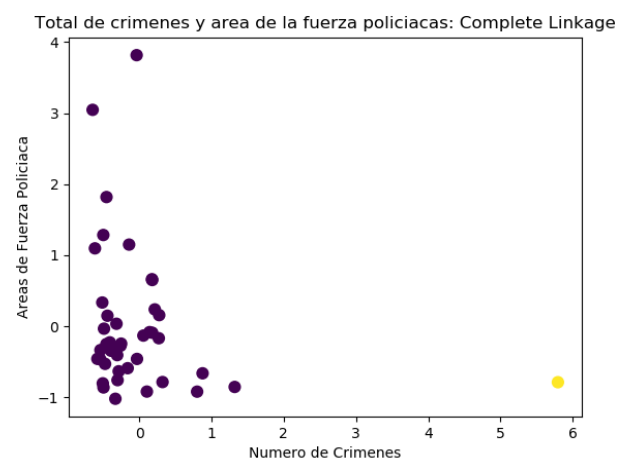
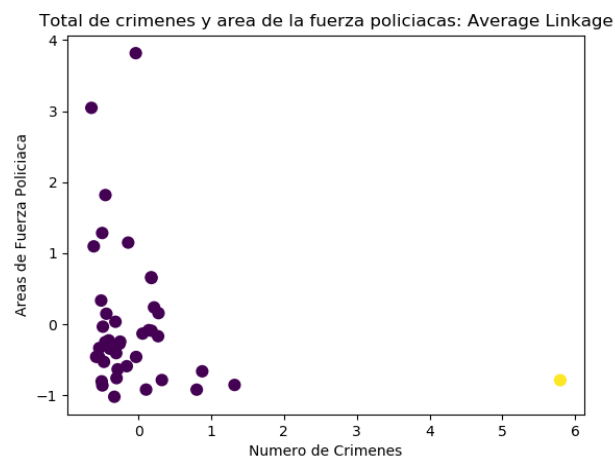
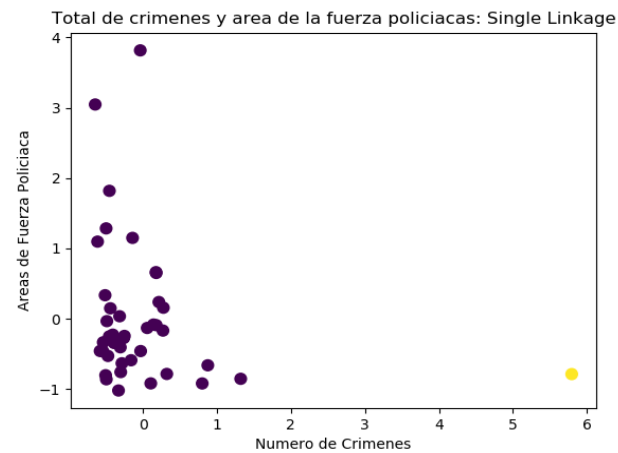
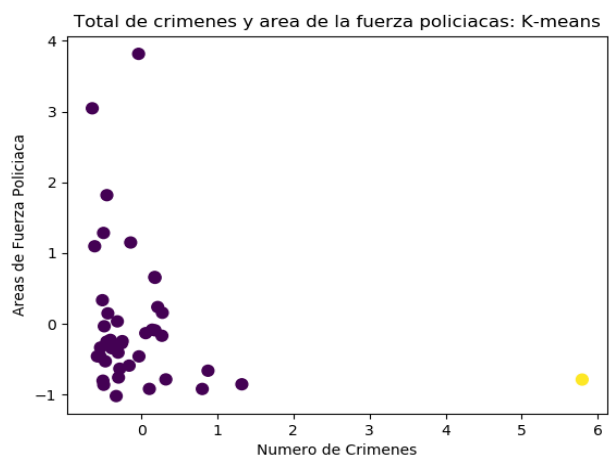
K=3 Silhouette\_score=0.347

Se presentaran los resultados del conjunto de datos de “Total de Crimenes y área de la fuerza policiaca”

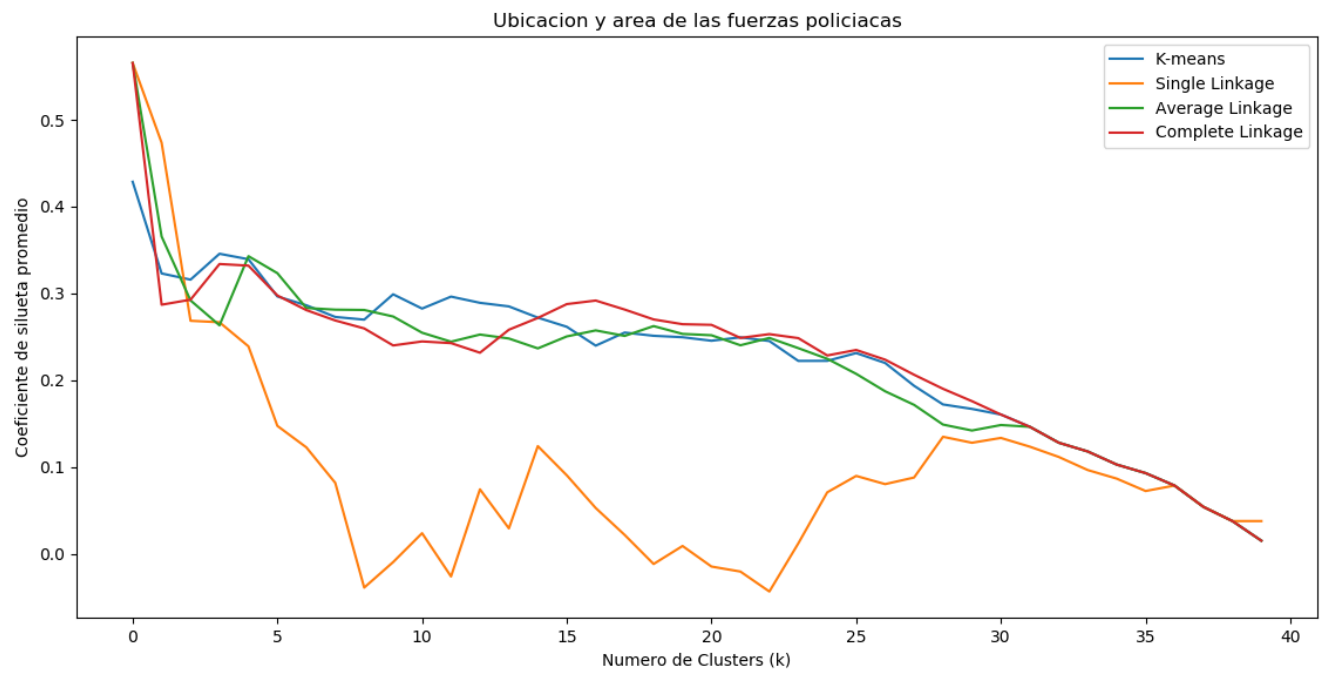


En esta grafica se puede ver que el coeficiente mas alto de los cuatro algoritmos aplicados no solo es el mismo, sino que coincide con el mismo numero de clusters (K) para cada uno. Tambien seguimos viendo como los coeficientes del algoritmo “single linkage” se siguen diferenciando a comparacion de los otros tres que son mas similares.

**Graficas de dispersión generadas con los valores de K de acuerdo al mejor coeficiente de cada algoritmo. (Conjunto de datos con 2 atributos)**

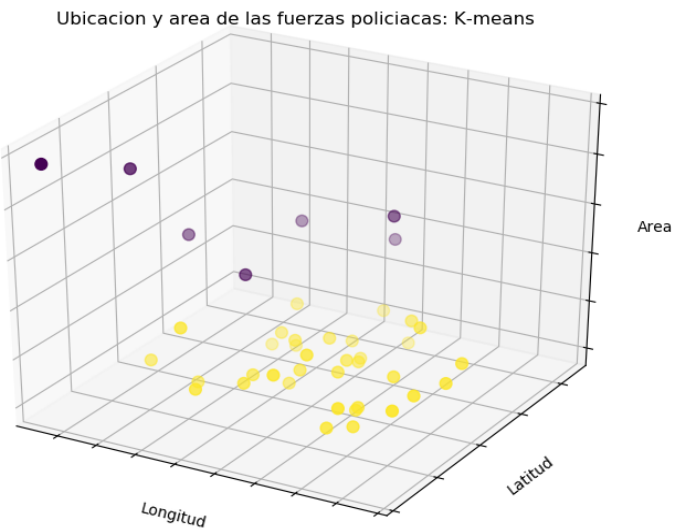


Se presentaran los resultados del conjunto de datos de “Ubicación y área de la fuerza policiaca”

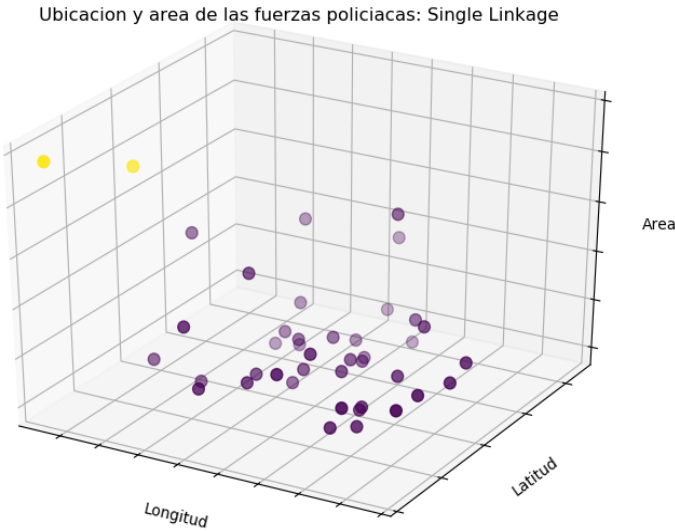


En la grafica se muestra que los coeficientes mas altos pertenecen al mismo numero de clusters (K) de los tres algoritmos aglomerativos. En cambio, el mejor de k-means está un poco por debajo de los demas. El coeficiente mas bajo de Single linkage es un poco mayor al de los otros que son el mismo.

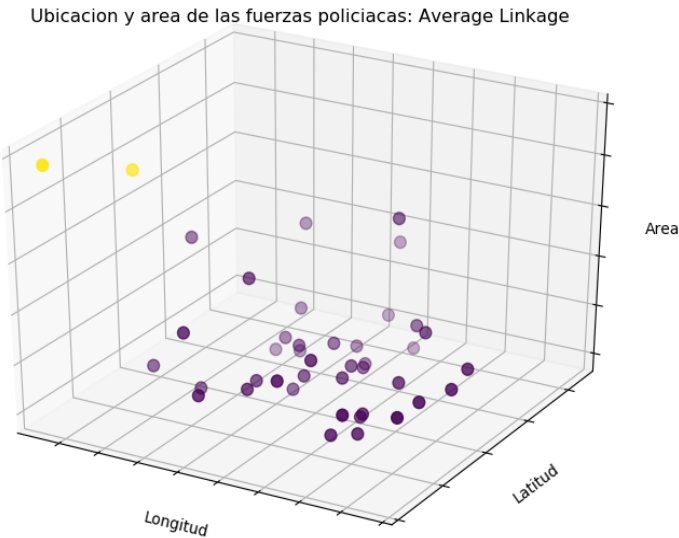
**Graficas de dispersión generadas con los valores de K de acuerdo al mejor coeficiente de cada algoritmo. (Conjunto de datos con 3 atributos)**



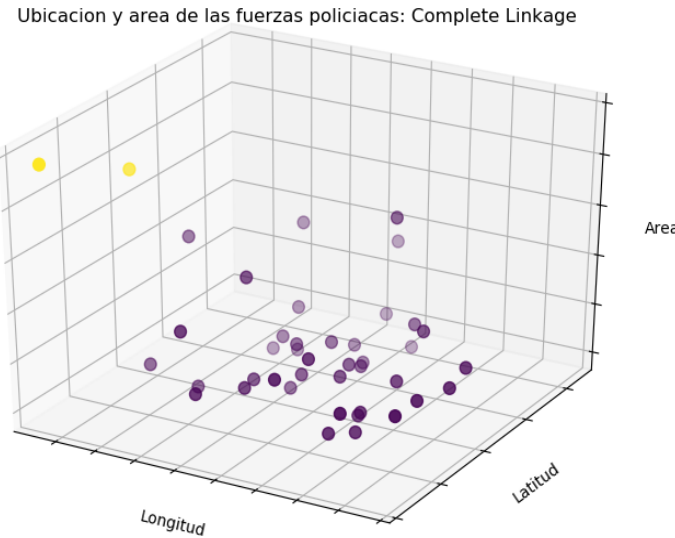
K=2 Silhouette\_score=0.428



K=2 Silhouette\_score=0.565

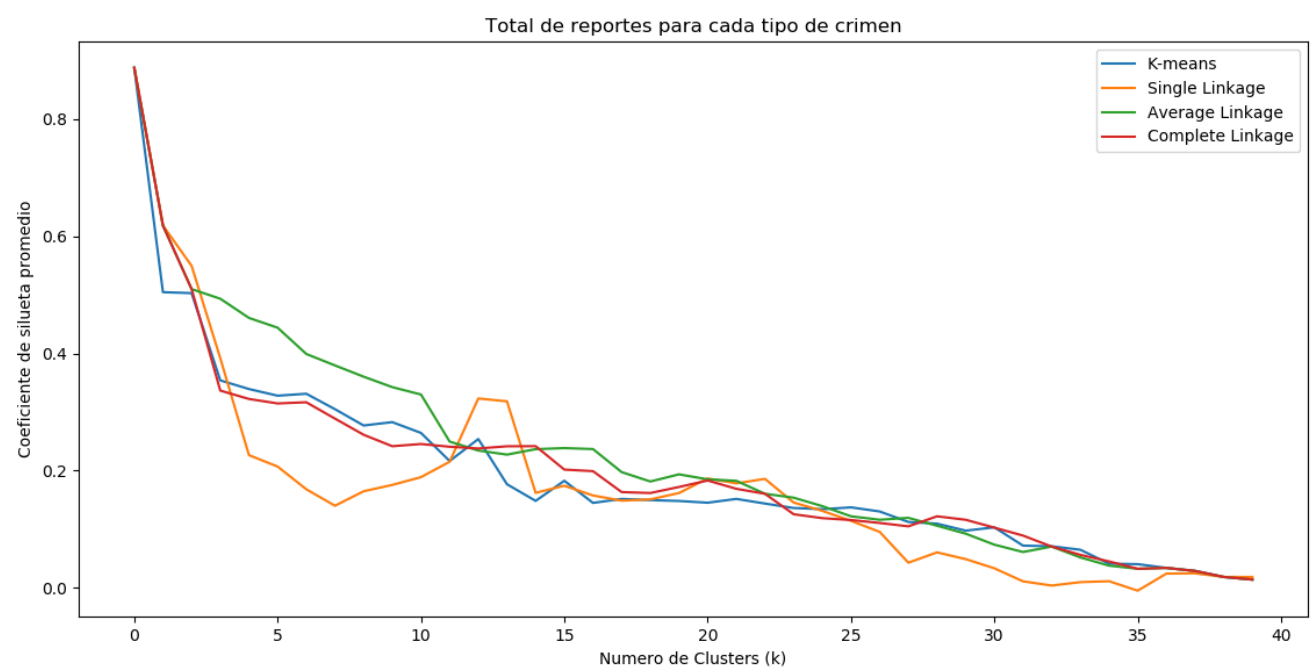


K=2 Silhouette\_score=0.565



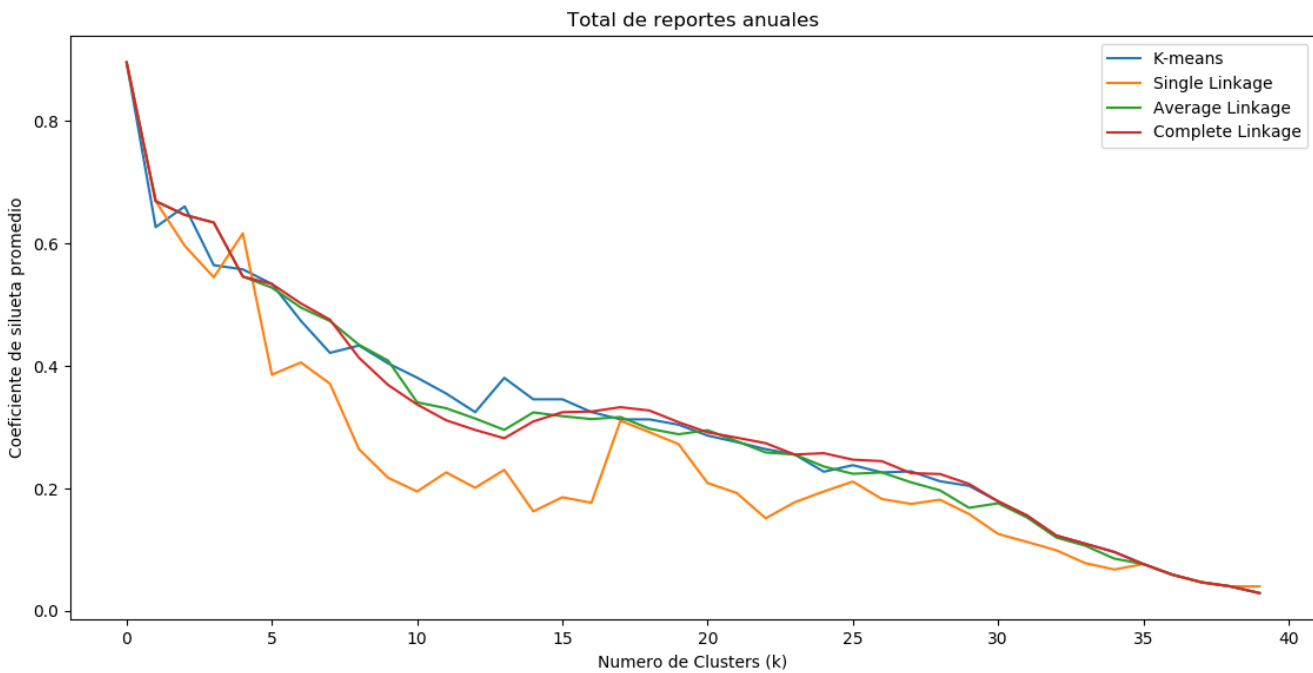
K=2 Silhouette\_score=0.565

Se presentaran los resultados del conjunto de datos de “Total de reportes por cada tipo de crimen”



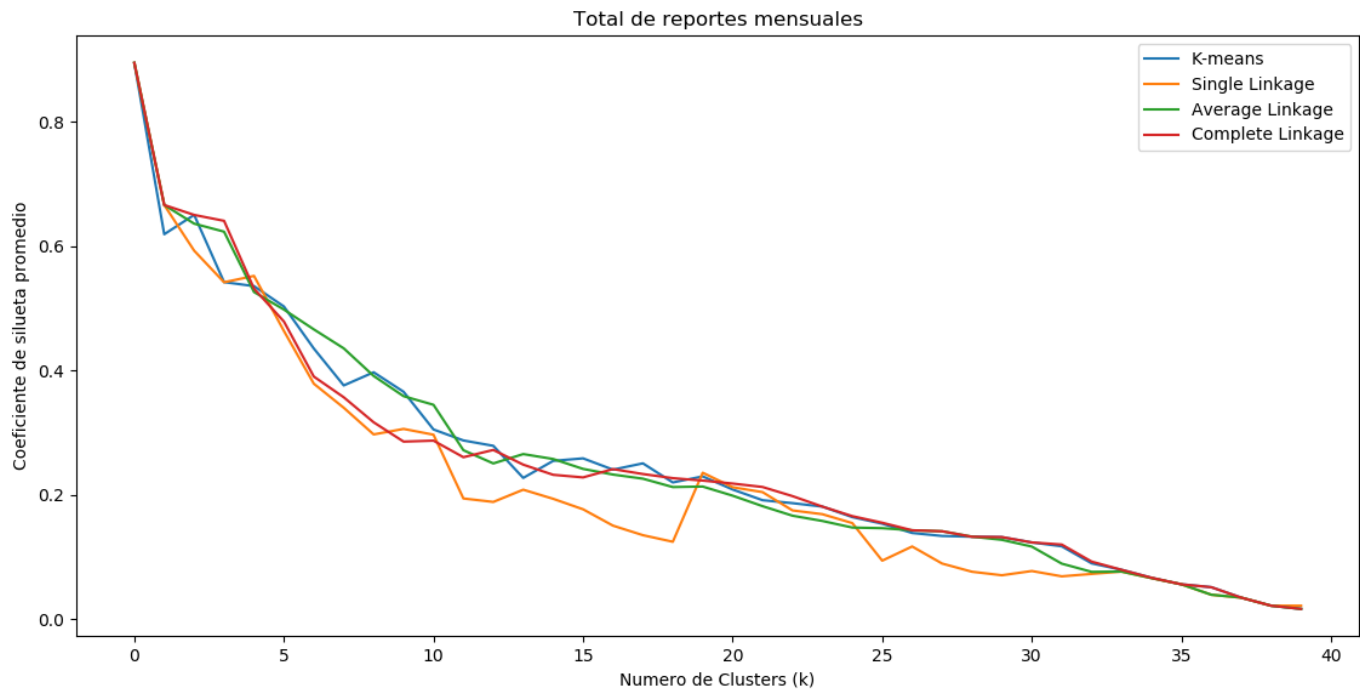
El coeficiente de silueta mas alto se aplica con 2 clusters (k) para cada algoritmo. Los coeficientes de los algoritmos de Kmeans, Average y Complete Linkage siempre disminuyen conforme a que aumenta K, en cambio, Single Linkage aumenta y disminuye cierto tiempo. El coeficiente mas bajo es casi el mismo para los cuatro algoritmos.

Se presentaran los resultados del conjunto de datos de “Total de reportes anuales”



El coeficiente de silueta mas alto se aplica con 2 clusters (k) para cada algoritmo. Los coeficientes de los algoritmos de Average y Complete Linkage casi siempre disminuyen conforme a que aumenta K, en cambio, Kmeans y Single Linkage aumentan y disminuyen cierto tiempo. El coeficiente mas bajo de Single linkage es un poco mayor al de los otros que son el mismo.

Se presentaran los resultados del conjunto de datos de “Total de reportes mensuales”



El coeficiente de silueta mas alto se aplica con 2 clusters (k) para cada algoritmo. Los coeficientes de los algoritmos de Average y Complete Linkage casi siempre disminuyen conforme a que aumenta K, en cambio, Kmeans y Single Linkage aumentan y disminuyen cierto tiempo. El coeficiente mas bajo es casi el mismo para los cuatro algoritmos.