



PRESENTACIÓN

BICIALPES

Daniel Esteban Aguilera Figueroa – 202010592

Laura Vanessa Martínez Prieto – 202012624

Cristian Armando Sánchez Ocampo – 202022112

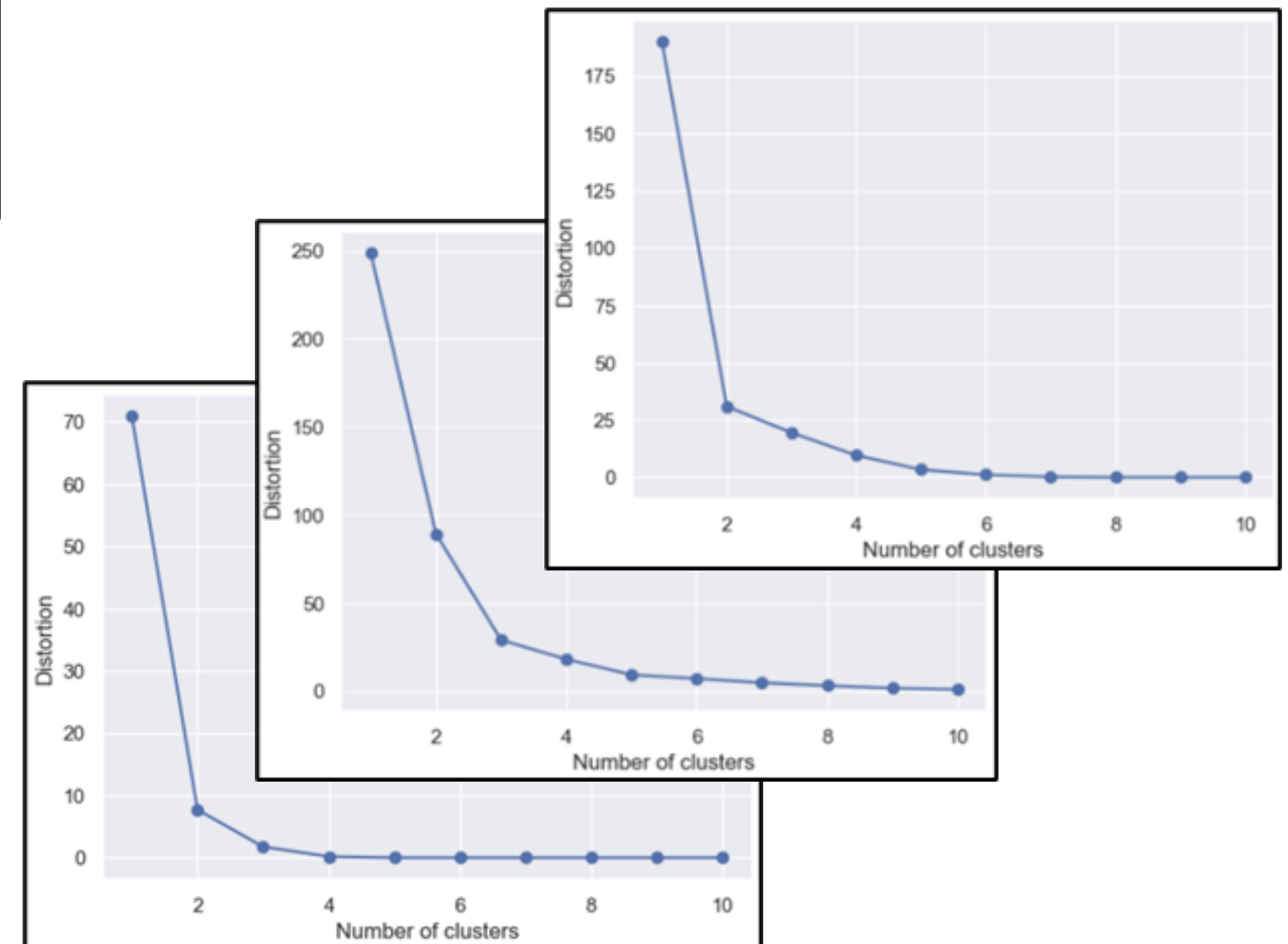
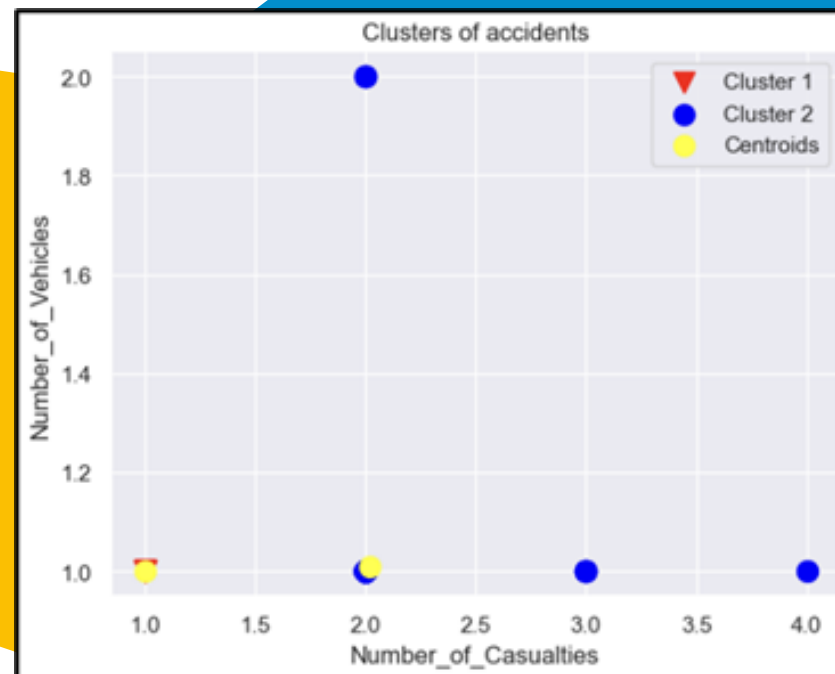
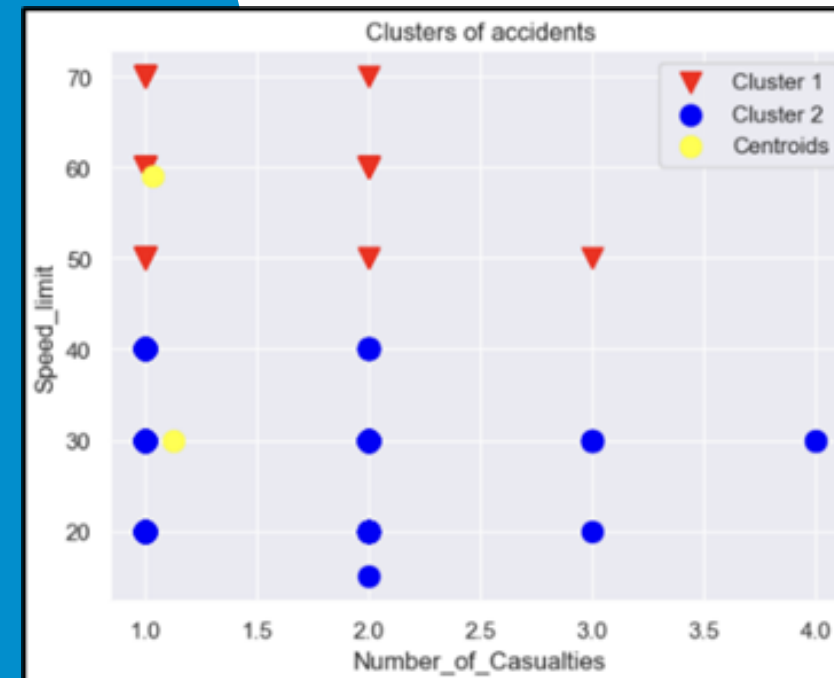
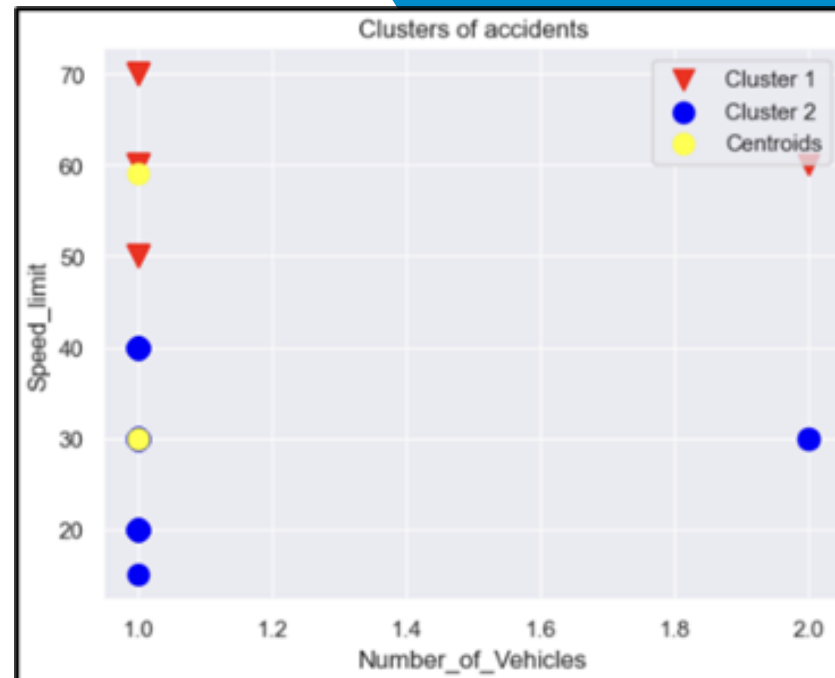
UNIVERSIDAD DE LOS ANDES

INTRODUCCIÓN

Se generó una caracterización e identificación de posibles patrones que pueden ser utilizados para entender los factores que indican a la severidad de los accidentes con el objetivo de trazar mapas de seguridad vial. Se espera aplicar tres diferentes algoritmos de clustering para lograr resolver el objetivo entregado por la organización.

K-MEANS

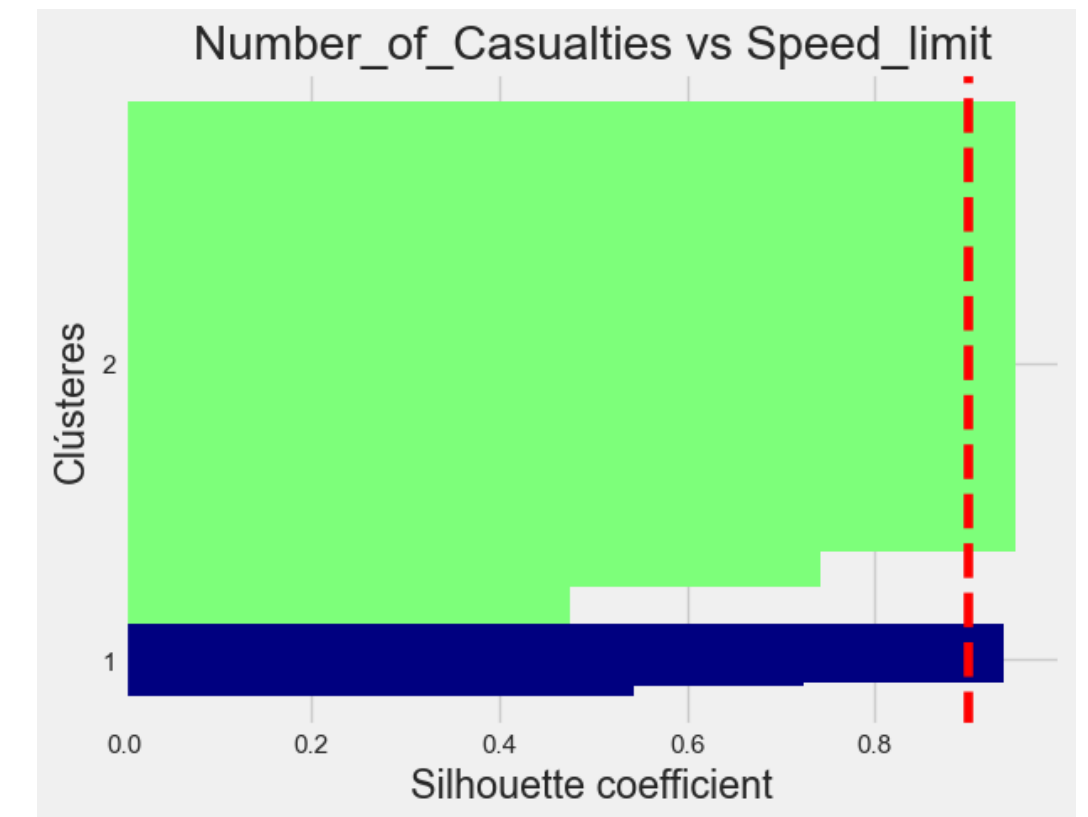
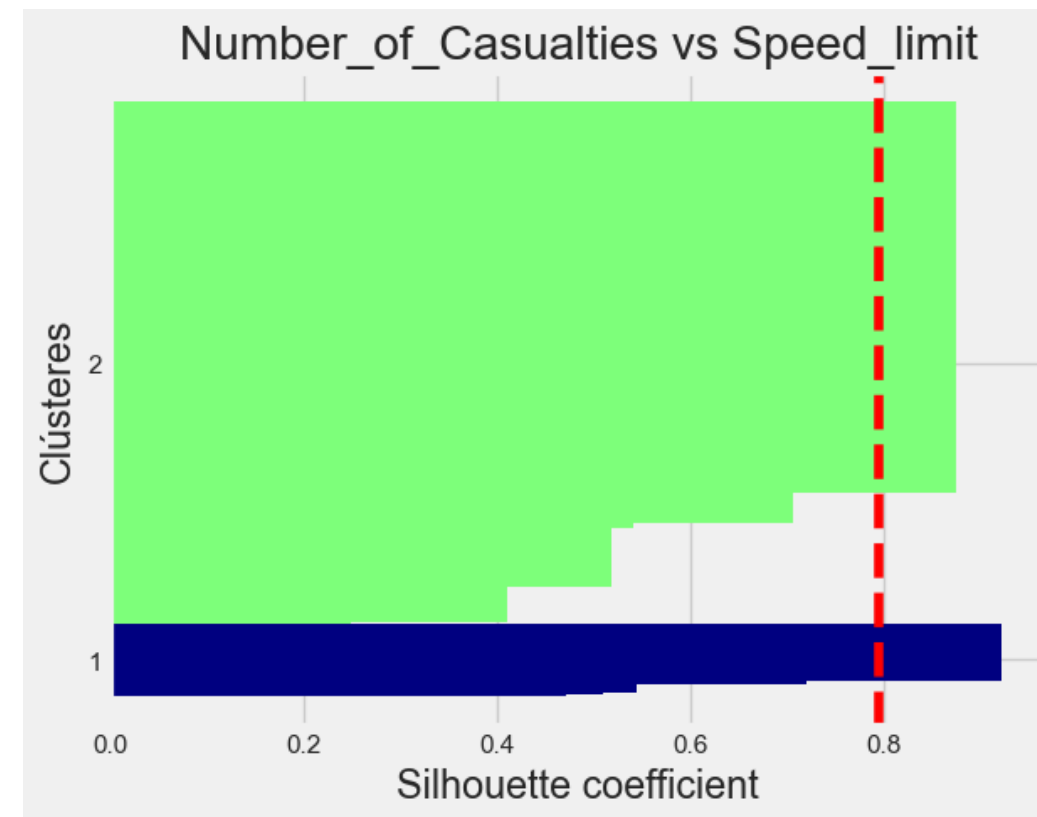
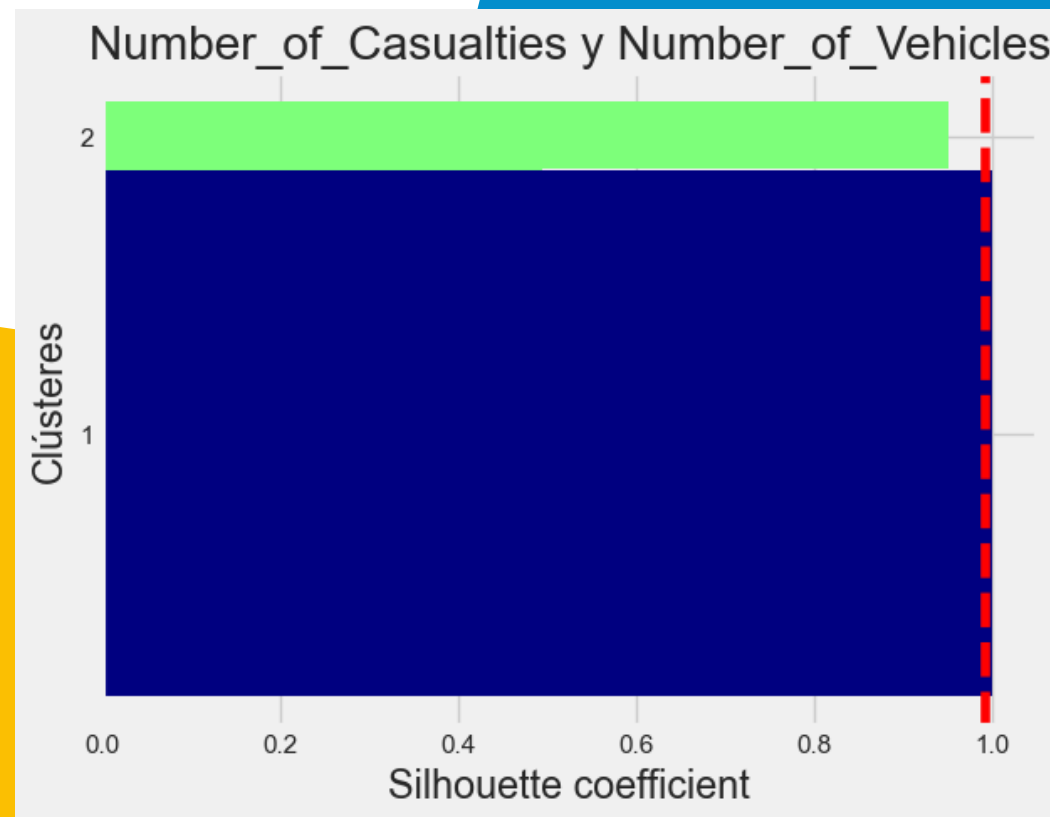
Este algoritmo en específico fue usado para analizar el comportamiento de las variables numéricas.



Se nota variabilidad en los datos con una fuerte tendencia a 1

Dos clusters

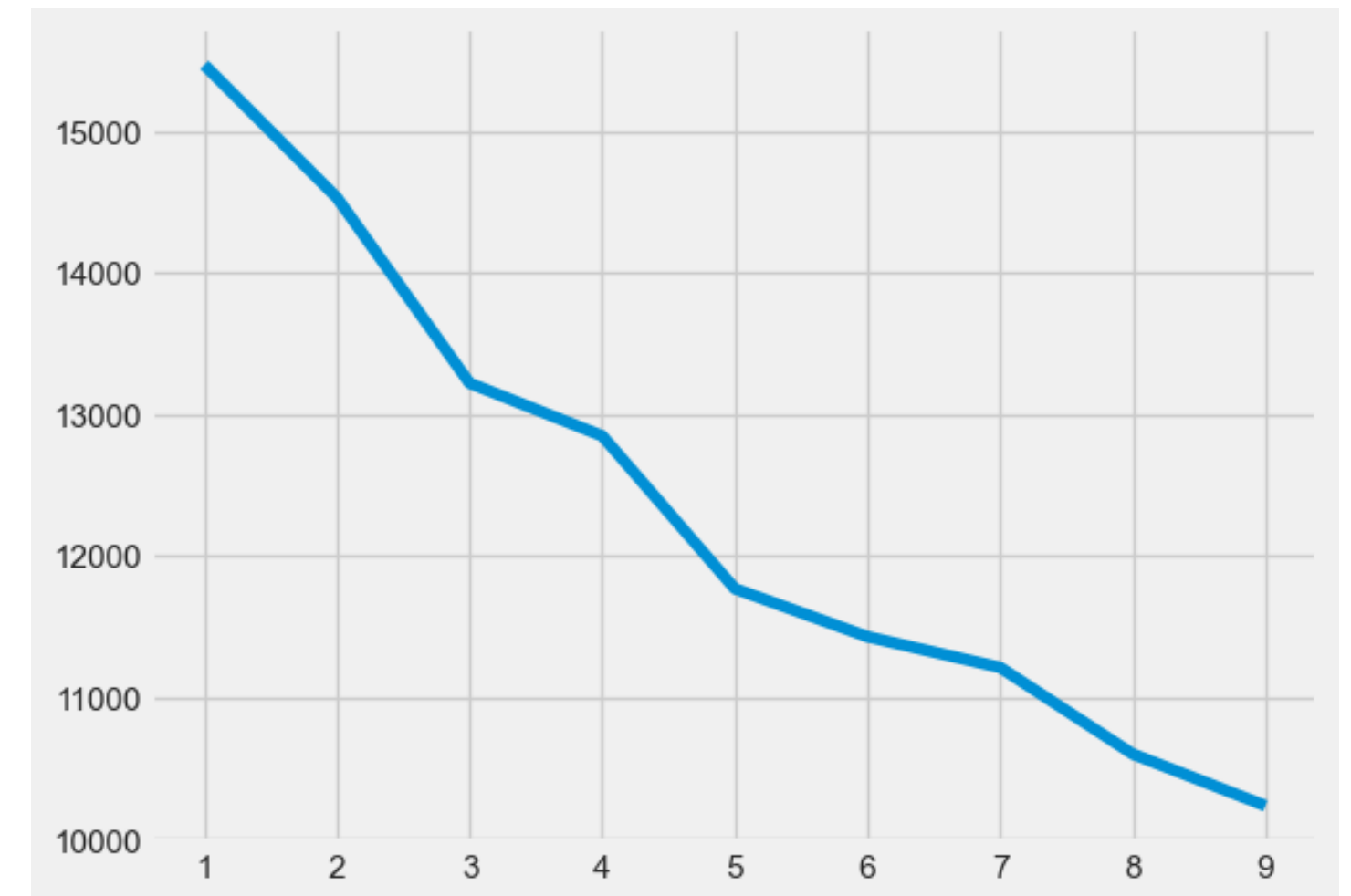
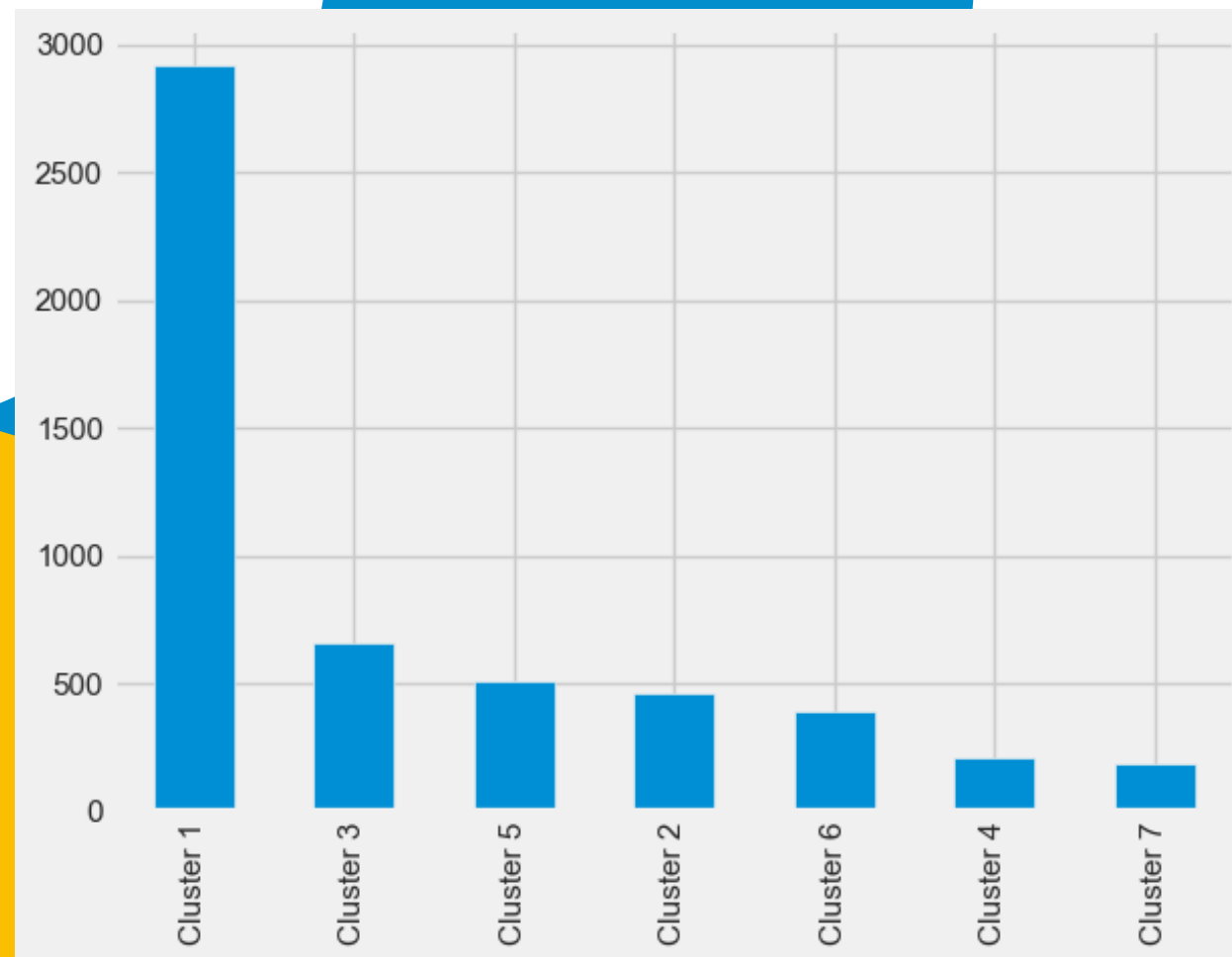
K-MEANS



Se nota la tendencia a 1 por parte de las siluetas indicando un buen modelo y variabilidad.

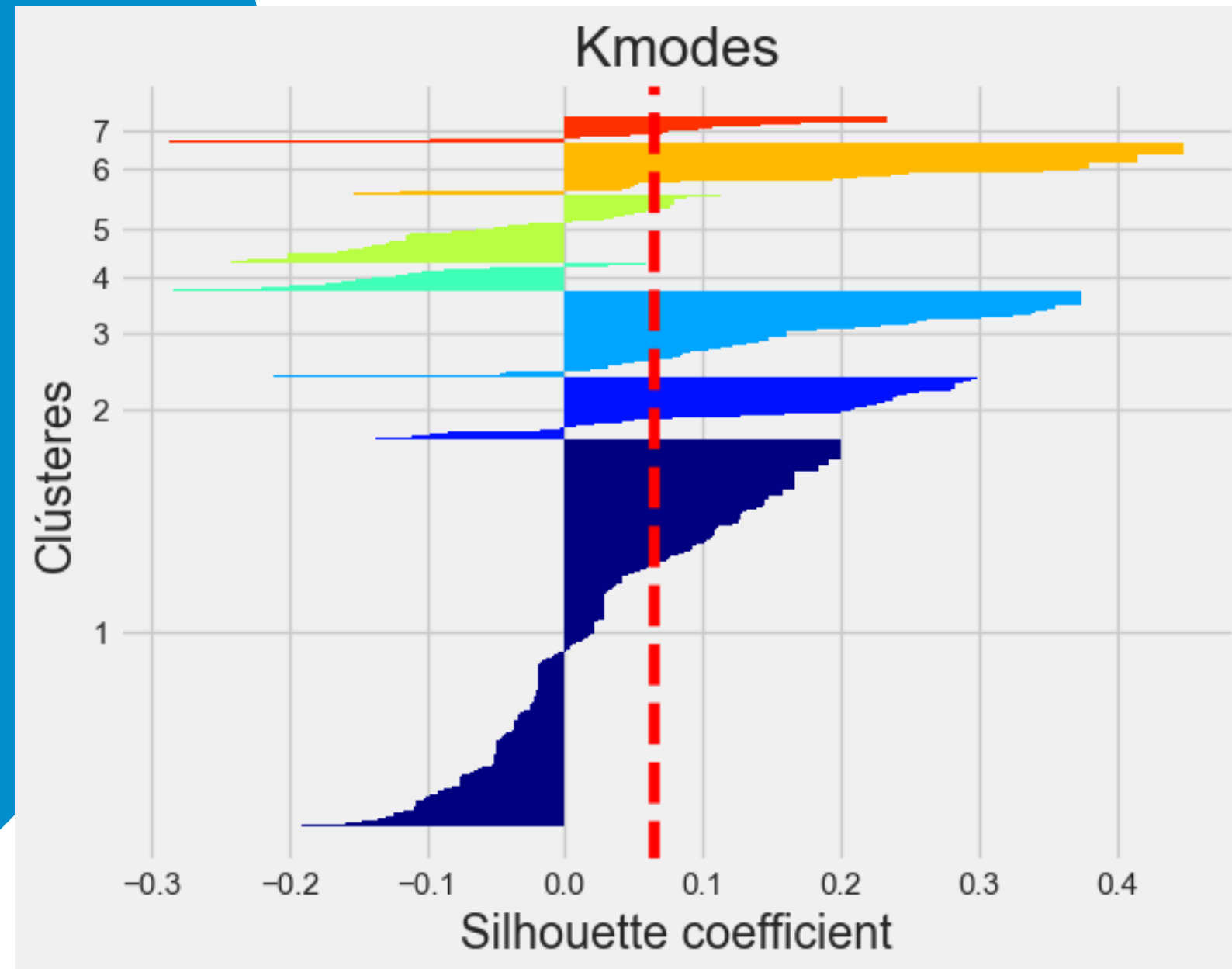
K-MODES

Este algoritmo en específico fue usado para analizar el comportamiento de las variables categóricas.



Se nota una gran tendencia de los datos a pertenecer en el cluster 1.

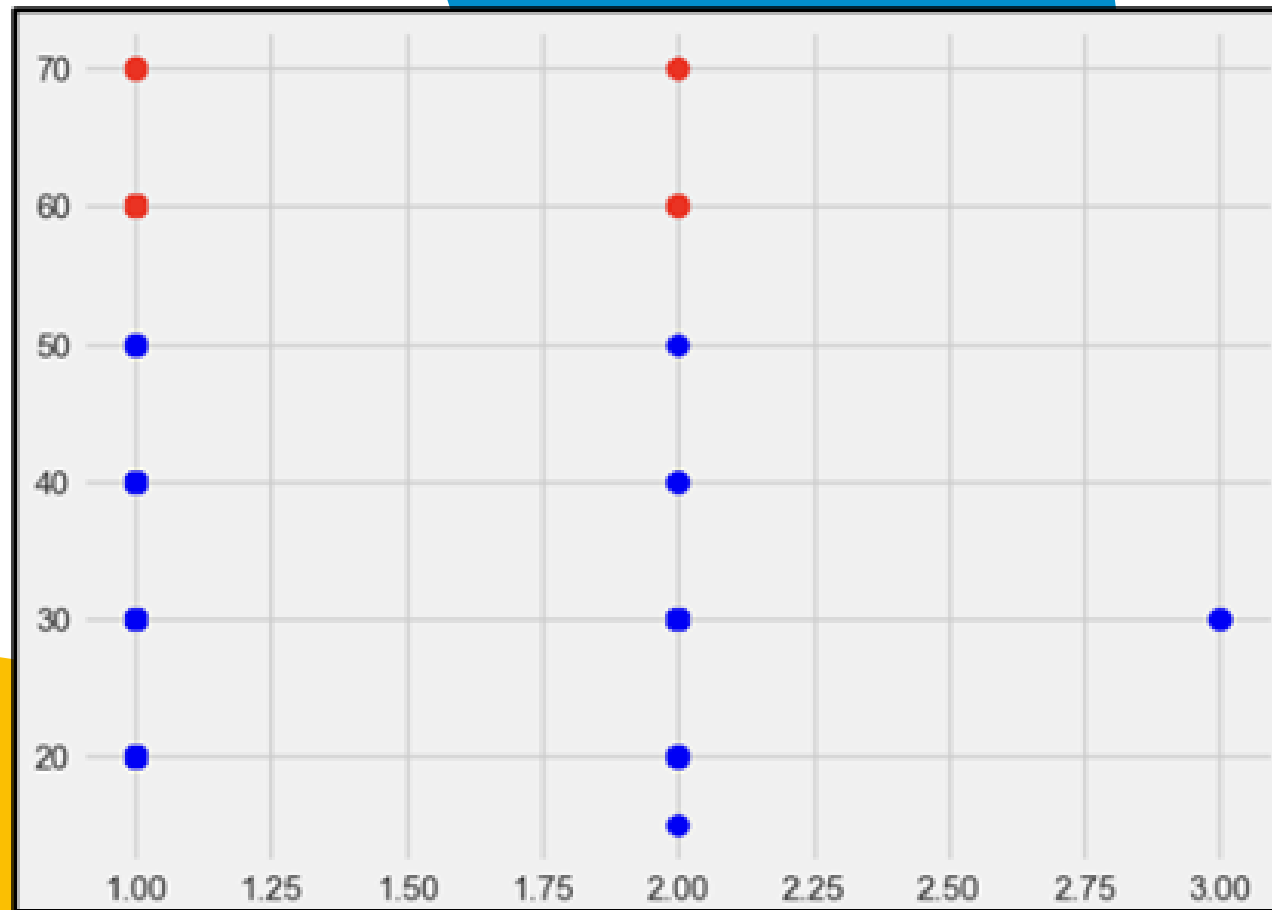
K-MODES



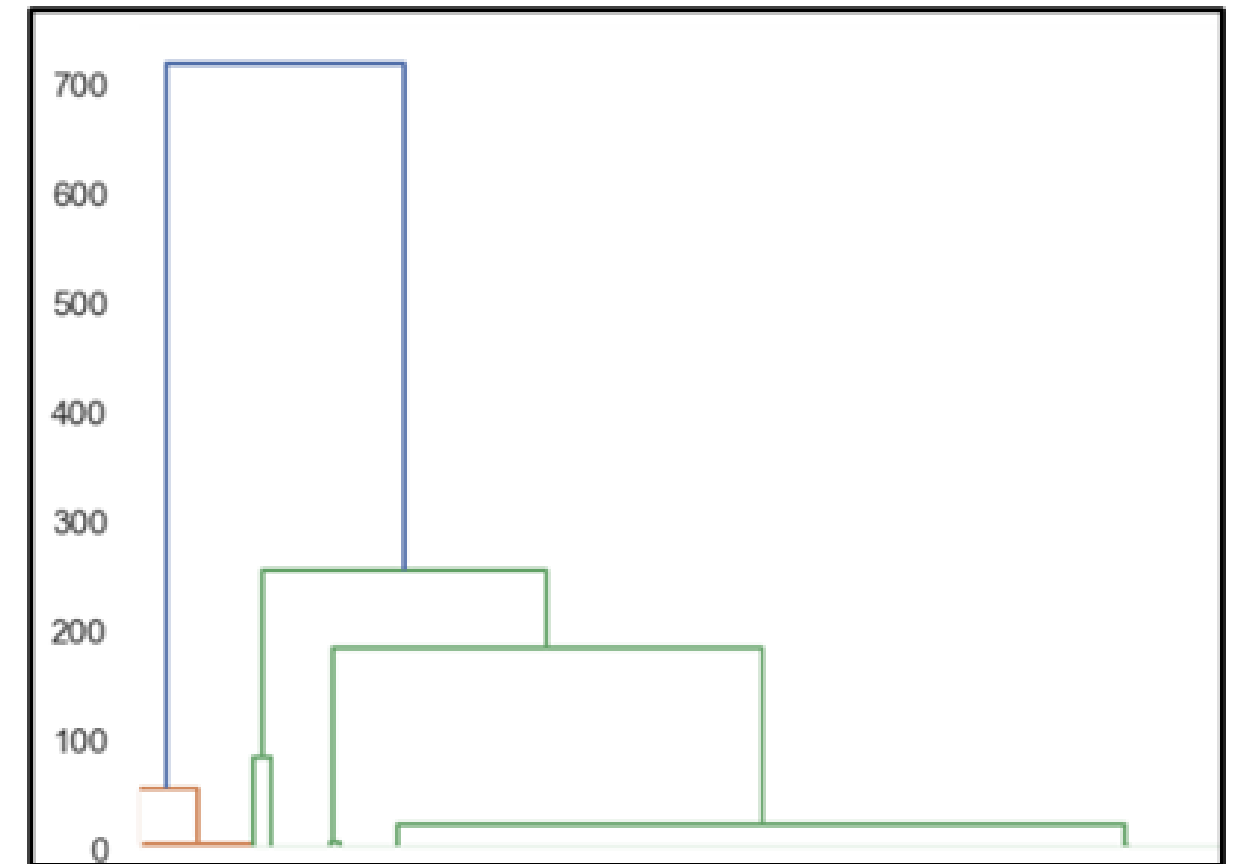
Se nota un comportamiento bastante variado en cuanto al puntaje de la silueta, indicando que algo sucedió con los clusters escogidos por el algoritmo para los datos.

AGLOMERATIVE CLUSTERING

Se decidió implementar este algoritmo en el caso de Speed_Limit vs Number_of_Casualties, para así lograr un mejor análisis y lograr una comparación con el gráfico entregado por el algoritmo de k-means.

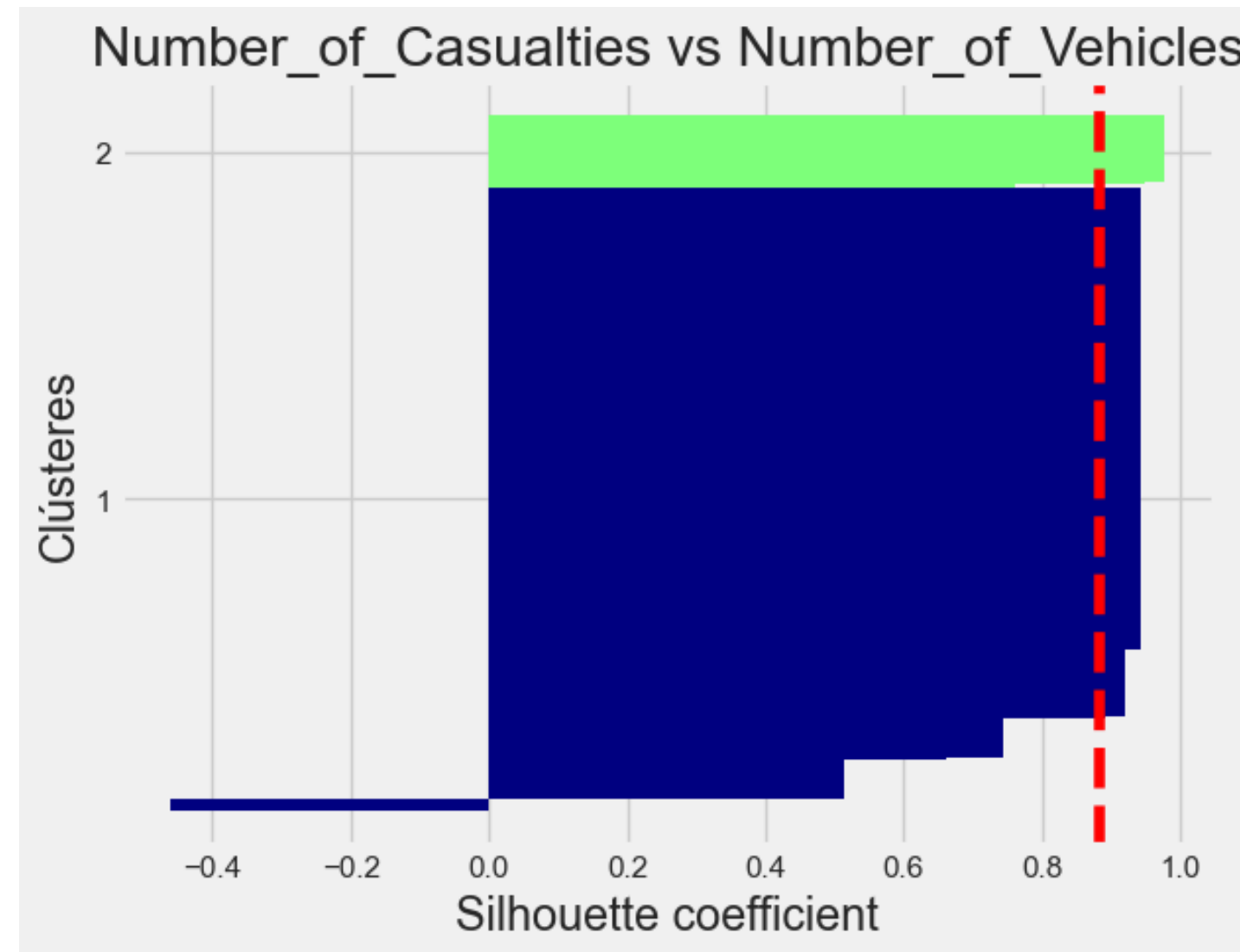


De igual forma que el generado por el algoritmo de k-means, este algoritmo entrega un gráfico en donde los datos son variados con respecto al eje de "Speed_limit" y datos que tienden a uno en el eje de "Number_of_Casualties".



Dos clusters

AGLOMERATIVE CLUSTERING



Se nota la tendencia a 1 por parte de las siluetas.
Y un fragmento negativo indicando fallos en las
variables o los datos asociados a estas



CONCLUSIONES

Al hacer el análisis de los algoritmos se puede concluir que, para las necesidades del negocio, el algoritmo K-modes permite un mejor análisis sobre los datos, ya que se tiene en cuenta las variables categóricas (factores que influyen en la accidentalidad). Por lo anterior, se puede saber que la mayoría de los accidentes son de bicicletas y que por lo general son leves, además que los principales factores presentes en la mayoría de accidentes.



Junction Detail=rotonda o minirrotonda

Weather Conditions=vientos suaves

Road Surface Conditions=condiciones secas

Light Conditions=luz diurna

Urban_or_Rural_Area=zonas urbanas

**Did_Police_Officer_Attend_Scene_of_Accident=
si se presentó un oficial de policía**

Day_of_Week=días laborales

Time=la tarde