**Instituto Tecnológico de Culiacán**



**Carrera: Ingeniería en Sistemas Computacionales**

**Materia: Temas Selectos de Bases de Datos**

**Alumnos:**

**José Alfredo García Aguilar**

**César Alfredo Astorga Ochoa**

**Trabajo: U3 T2 Minería de Datos**

**Fecha: 14-Mayo-2022**

**Horario de clase: 05:00 – 06:00 pm**

**Profesor: Daniel Esparza Soto**

* Vista utilizada en SQL Server:

create view vw\_accessoriescomponents as

select c.customerkey, c.geographykey, c.customeralternatekey, c.title, c.firstname, c.middlename, c.lastname, c.namestyle, c.birthdate, c.maritalstatus, c.suffix, c.gender, c.emailaddress, c.yearlyincome, c.totalchildren,

c.numberchildrenathome, c.englisheducation, c.spanisheducation, c.frencheducation, c.englishoccupation, c.spanishoccupation, c.frenchoccupation, c.houseownerflag, c.numbercarsowned, c.addressline1,

c.addressline2, c.phone, c.datefirstpurchase, c.commutedistance, g.city, g.stateprovincename, g.spanishcountryregionname, x.region, x.age, case x.[accessories] when 0 then 0 else 1 end as accessoriesbuyer, case x.[components] when 0 then 0 else 1 end as componentsbuyer

from dbo.dimcustomer as c

inner join dimgeography g on c.geographykey = g.geographykey

inner join(select customerkey, region, age, sum(case [englishproductcategoryname] when 'accessories' then 1 else 0 end) as accessories, sum(case [englishproductcategoryname] when 'components' then 1 else 0 end) as components

from dbo.vdmprep

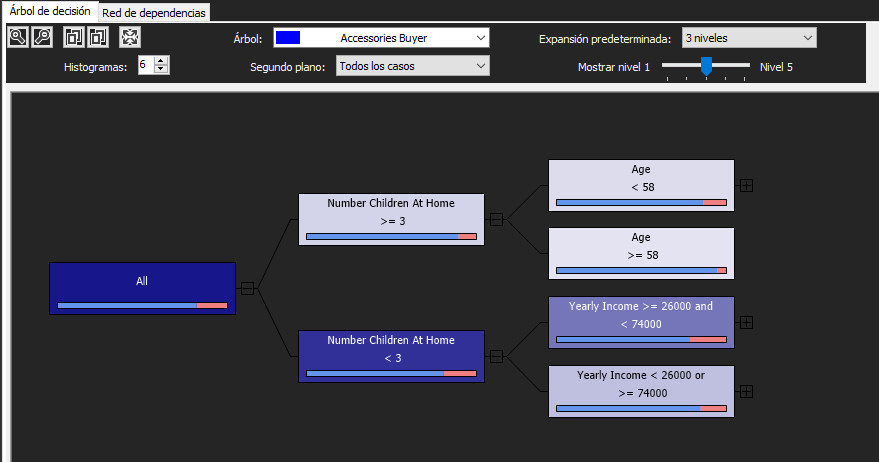
group by customerkey, region, age) as x on c.customerkey = x.customerkey

* Gráficas resultantes de los 3 algoritmos:

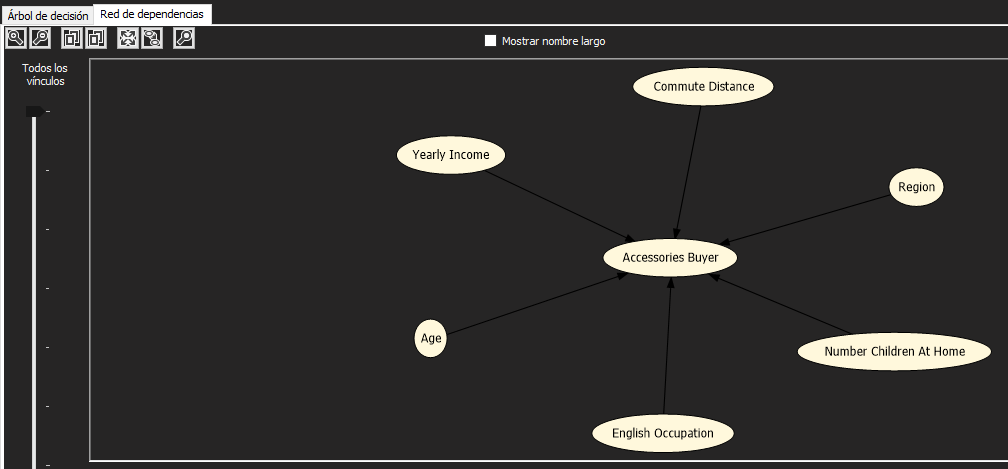
**Accesories Buyer:**

TM\_Decision\_Tree:

Árbol de decisión:

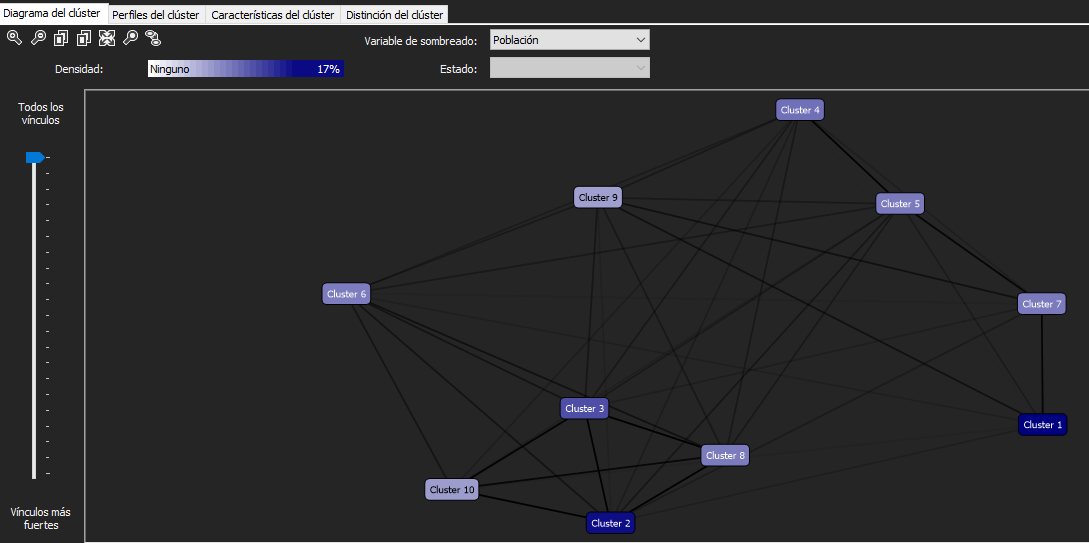


Red de dependencias:



TM\_Clustering:

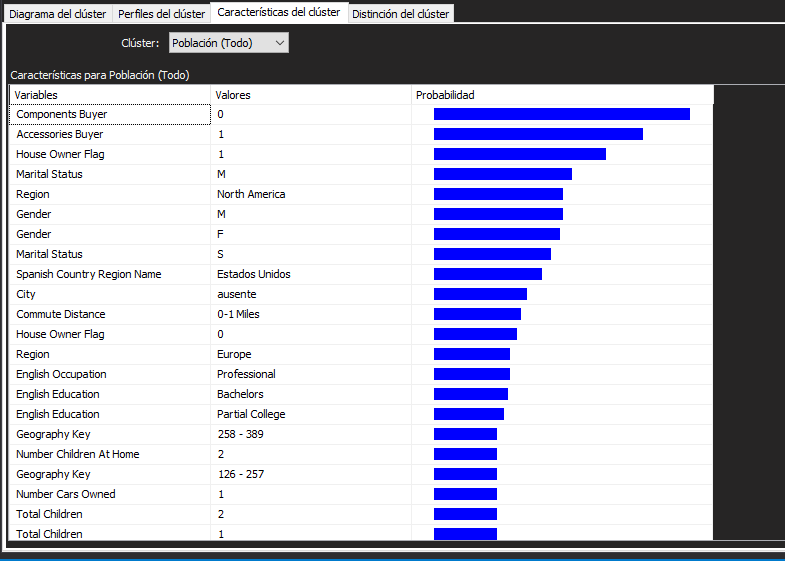
Diagrama de clúster:



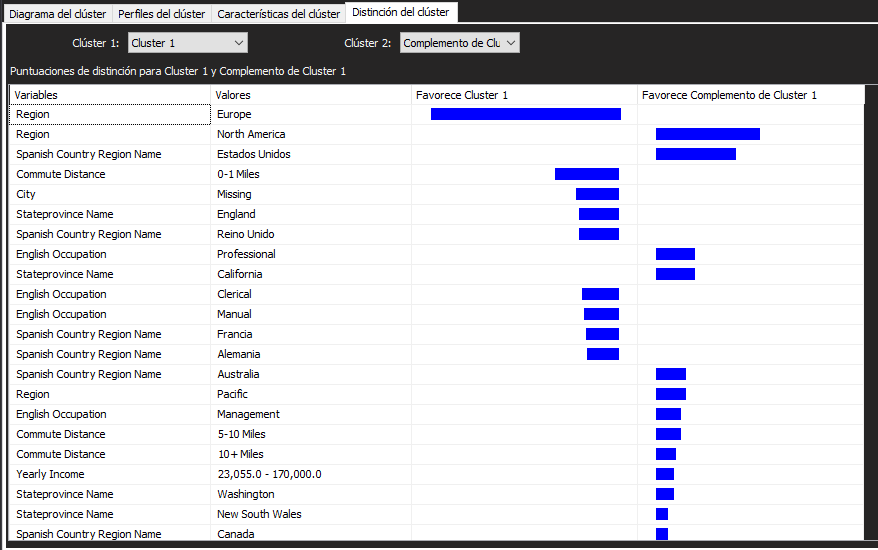
Perfiles de clúster:



Características del clúster:

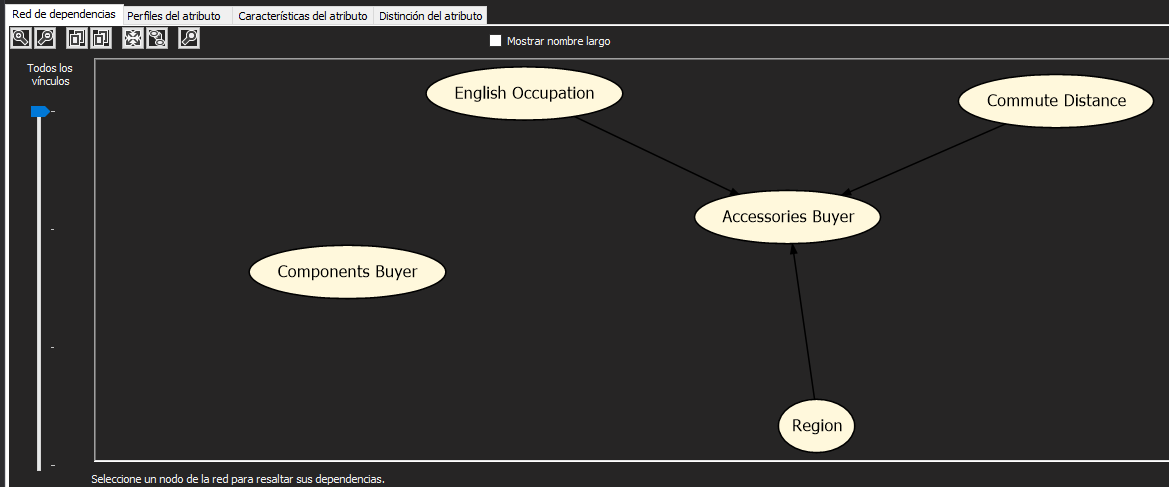


Distinción del clúster:

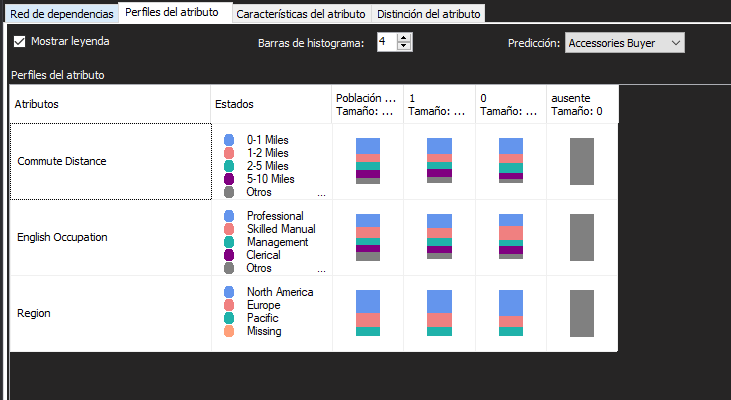


TM\_NaiveBayes:

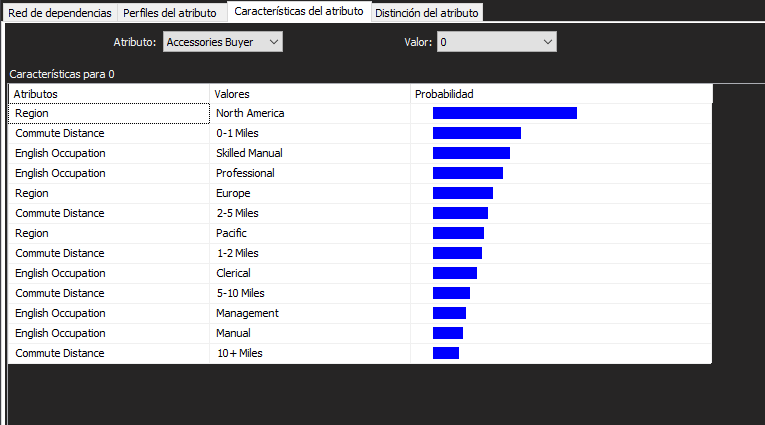
Red de dependencias:



Perfiles del atributo:



Características del atributo:

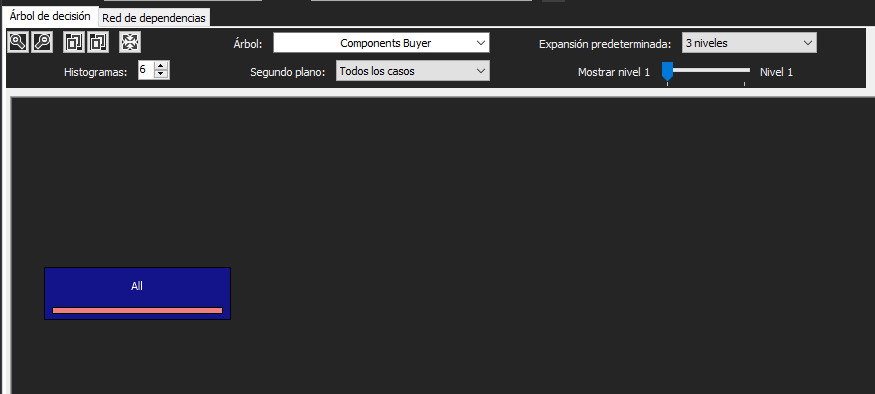


Distinción del atributo:

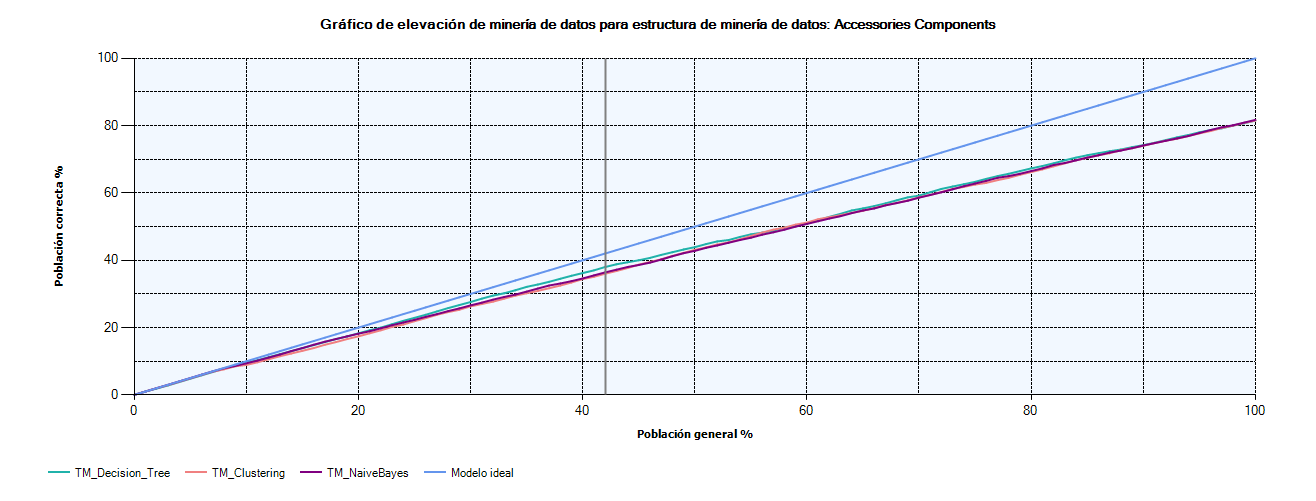


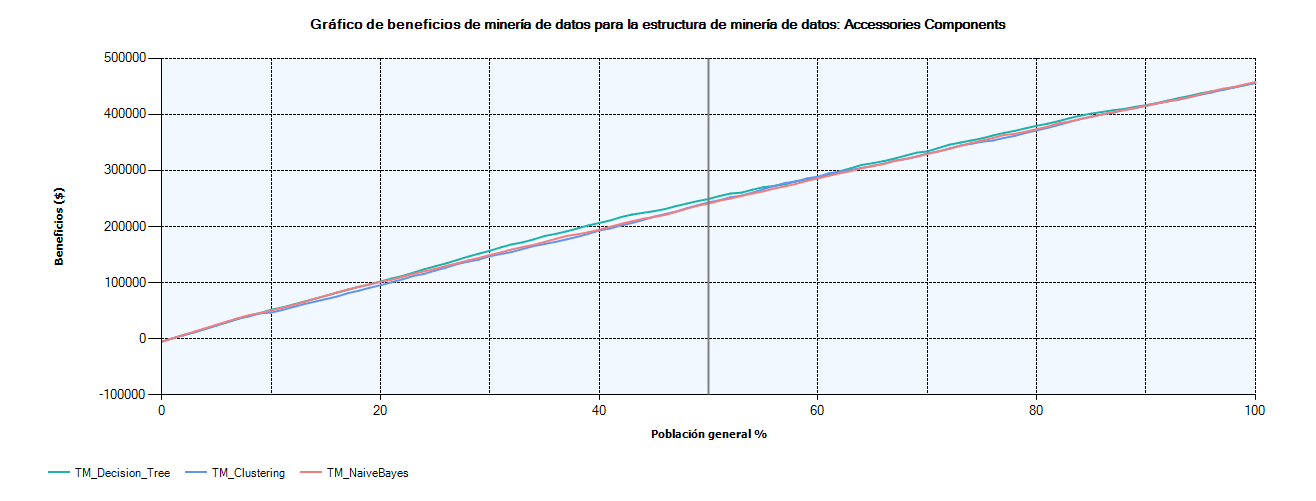
**Components Buyer:**

Árbol de decisión:



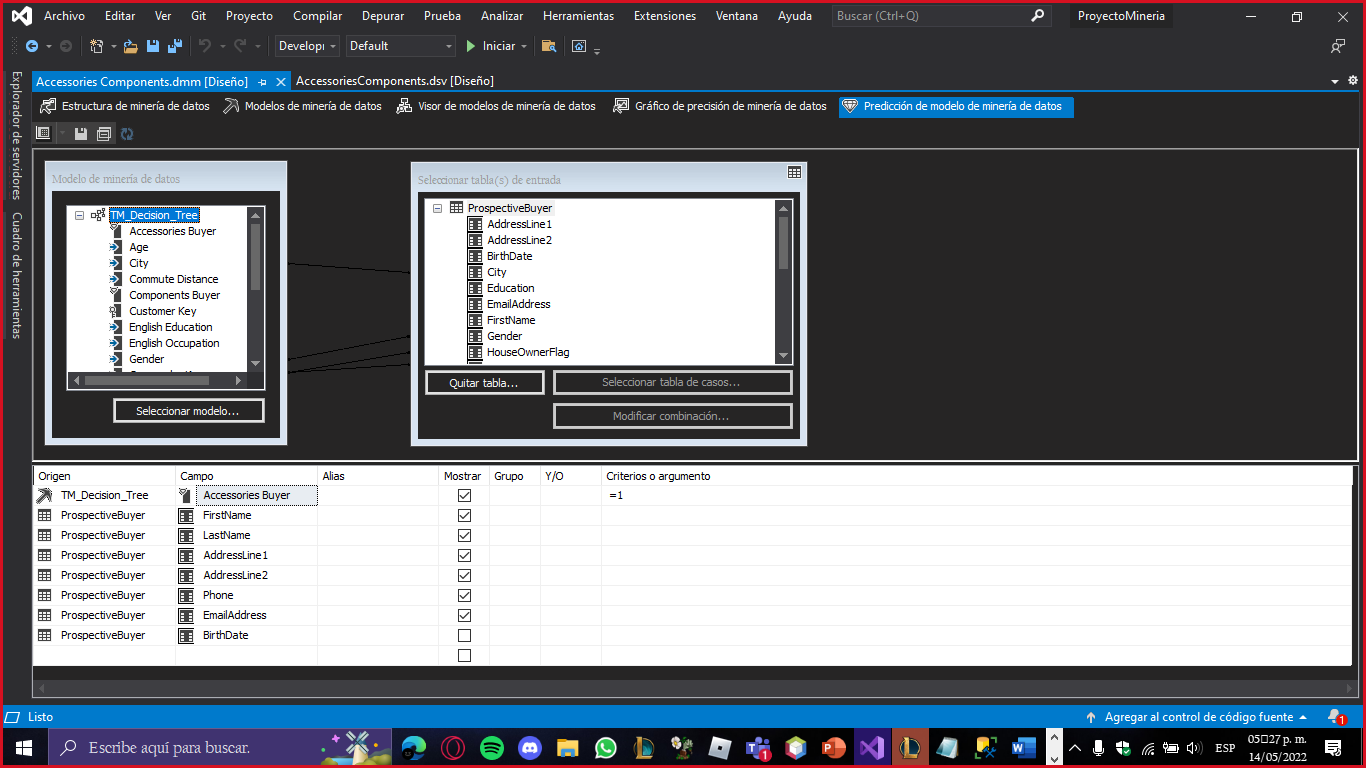
Nota: El components buyer sale vacio

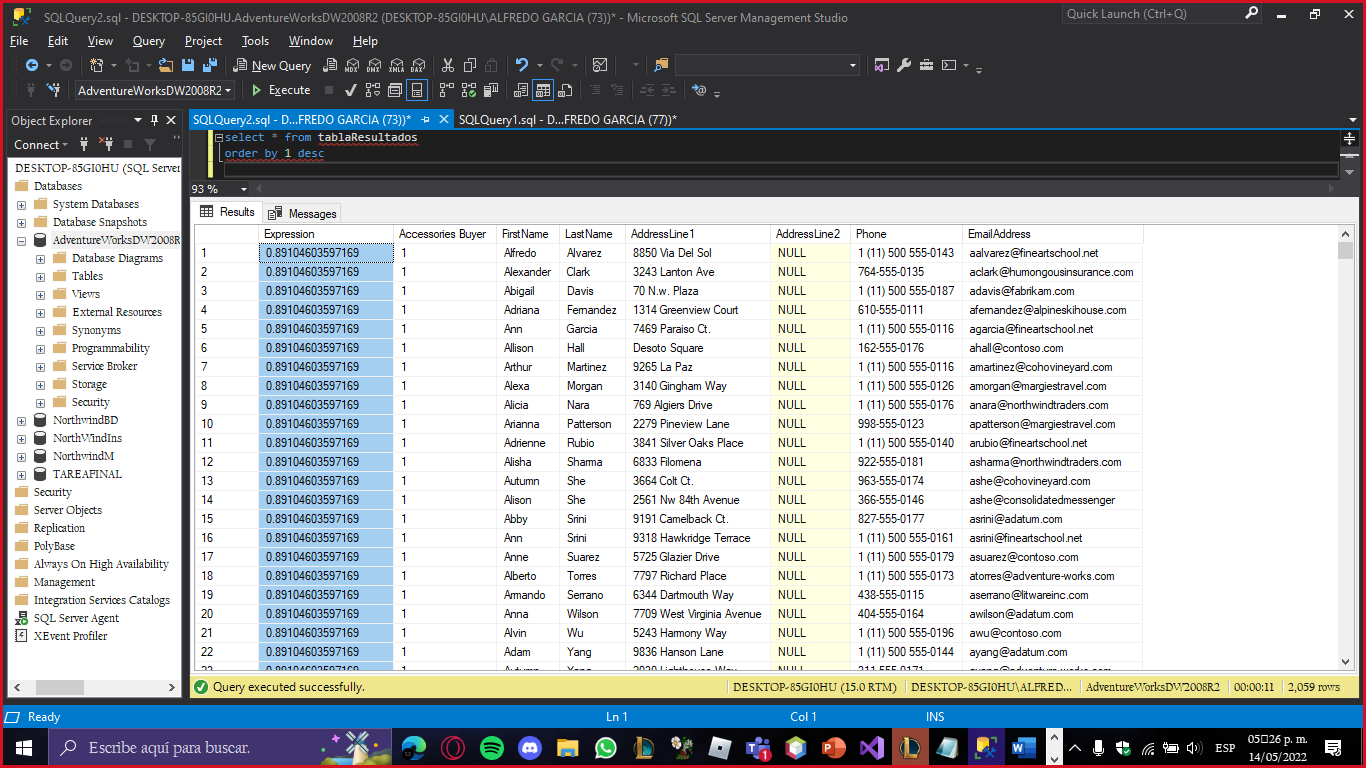
* Grafica de precisión evaluando los 3 algoritmos. Definir cuál es el algoritmo con mayor precisión.



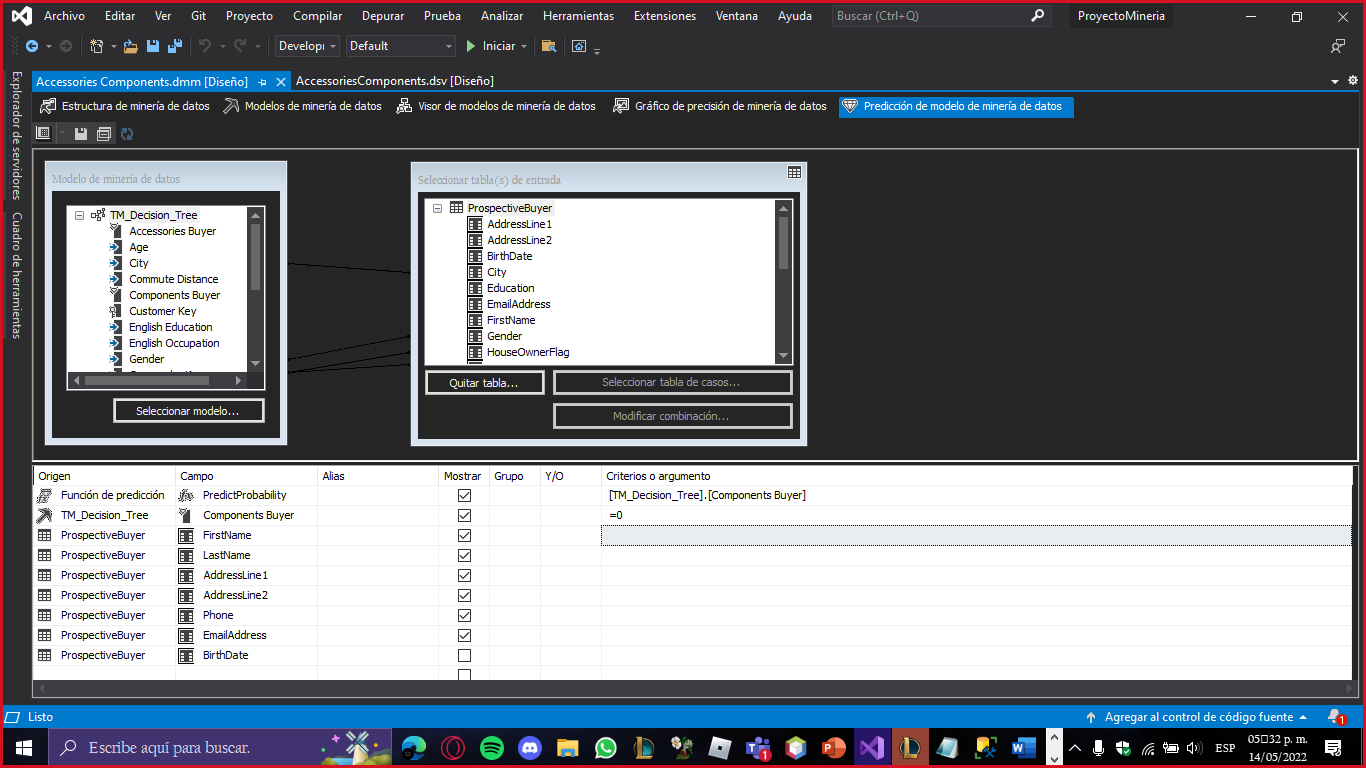
El árbol de decisión (TM\_Decision\_Tree) es el algoritmo con mayor precisión porque es el que más se acerca a la línea ideal en ambas gráficas.

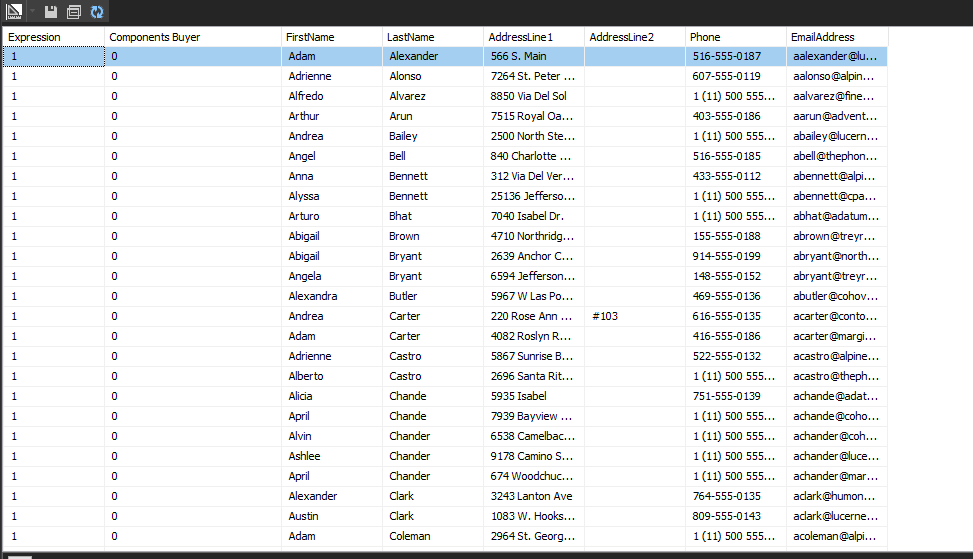
* Predecir nuevos clientes con la tabla ProspectiveBuyer, definir que algoritmo se utilizará para la predicción y mostrar los resultados

Accessories Buyer:



Components Buyer:





El árbol de decisión (TM\_Decision\_Tree) es el algoritmo a usar en este caso por tener mayor precisión.