|  |  |
| --- | --- |
|  |  |
| Project 3 |  |
|  |  |
|  | Data Mining |
|  | Ahmed Salah Hammad120201578 |

**Title of Dataset:**

-Airline Passenger Satisfaction

**Description of Dataset:**

-This dataset contains an airline passenger satisfaction survey. What factors are highly correlated to a satisfied (or dissatisfied) passenger.

**Objective of Data Mining project:**

- Search for interesting relationship among items in Selected Attributes in the dataset.

- Predict the passengers who are likely to be disloyal Customer to make them Loyal Customers.

- Clustering Customers into distinct groups to find similarity among customers.

**Name of Attributes:**

**-** **Customer Type:** The customer type (Loyal customer, disloyal customer).

**-Age:** The actual age of the passengers

**-Type of Travel:**Purpose of the flight of the passengers (Personal Travel, Business Travel).

**-Class:** Travel class in the plane of the passengers (Business, Eco, Eco Plus)

**-Flight distance:** The flight distance of this journey

**-Inflight Wi-Fi service**: Satisfaction level of the inflight Wi-Fi service (0:Not Applicable;1-5)

**-Ease of Online booking:** Satisfaction level of online booking

**-Food and drink:** Satisfaction level of Food and drink

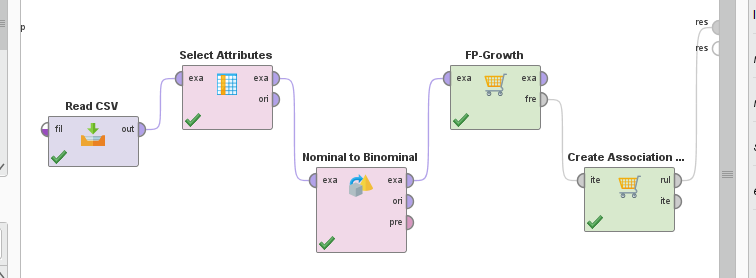
**-Satisfaction:** Airline satisfaction level (Satisfaction, neutral or dissatisfaction).

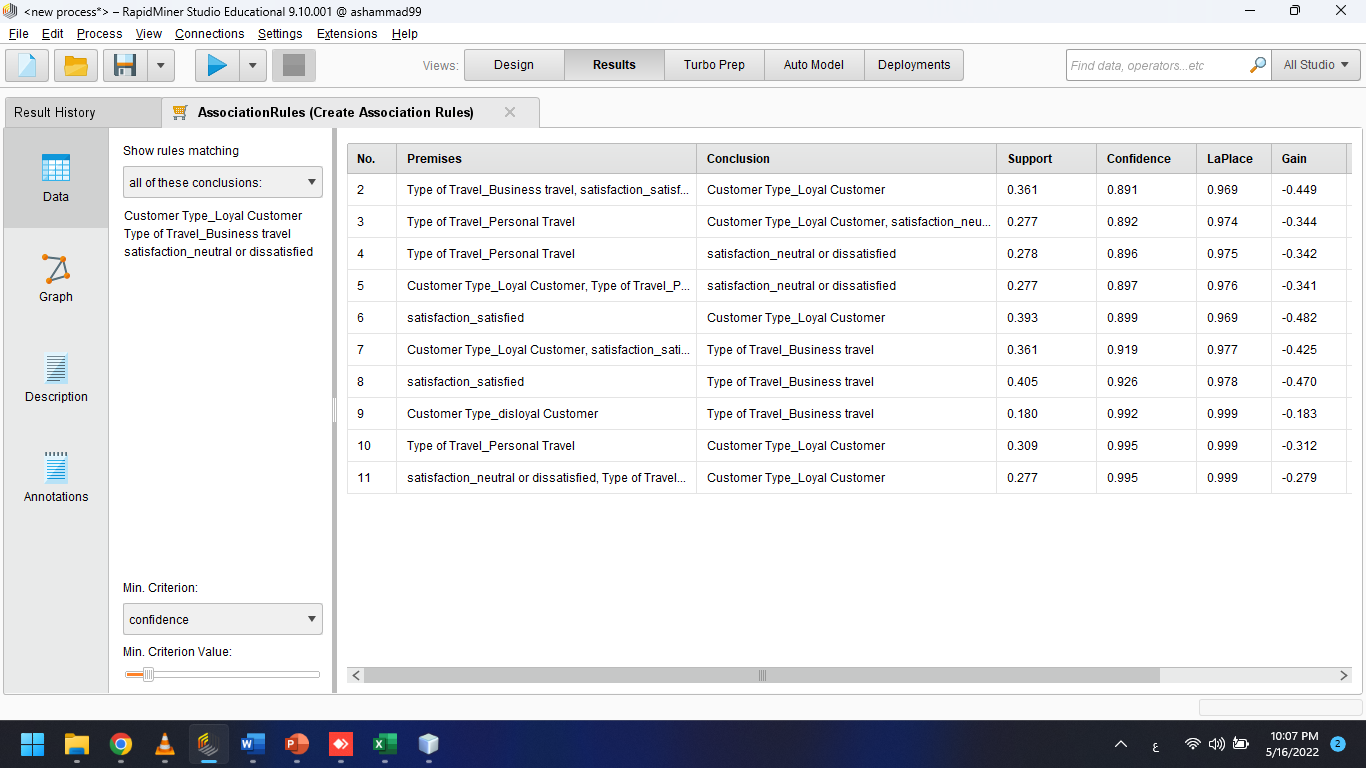
**Note:** the dataset has been cleaned by the owner, and I’ve deleted some rows because the large number in original dataset.

1. **Associations Rules:**

**Set Parameters of Operators:**

* **Select Attribute:** select 3 Attributes (customer type, satisfactions, travel type)
* **Nominal to Binominal:** 
  + Include special attributes
  + Transform Binominal
  + Use underscore in name
* **FP-Growth:** set min support = 0.7
* **Create Association Rules:** set min Confidence = 0.8

****Screenshot of process:

**Associations Rules screenshot:**

Explanations of 3 rules:

- if (type of travel=” Personal”) then (customer type=’ Loyal Customer’)

لو نوع رحلة السفر هو شخصي فانه من المحتمل ان يكون الزبون من المخلصين للشركة.

- if (stratification=’ satisfied’) then (type of travel=’ Business’)

لو الزبون كان راضي عن خدمات الشركة فمن المحتمل ان تكون نوع الرحلة خاصة بالبزنس.

- if (stratification=’ satisfied’) then (customer type=’ Loyal Customer’)

لو الزبون كان راضي عن خدمات الشركة فمن المحتمل ان يكون الزبون من المخلصين للشركة.

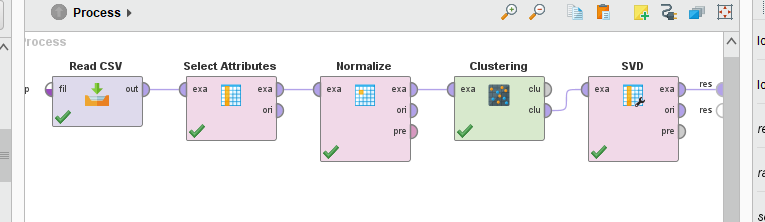
1. **Clustering:**

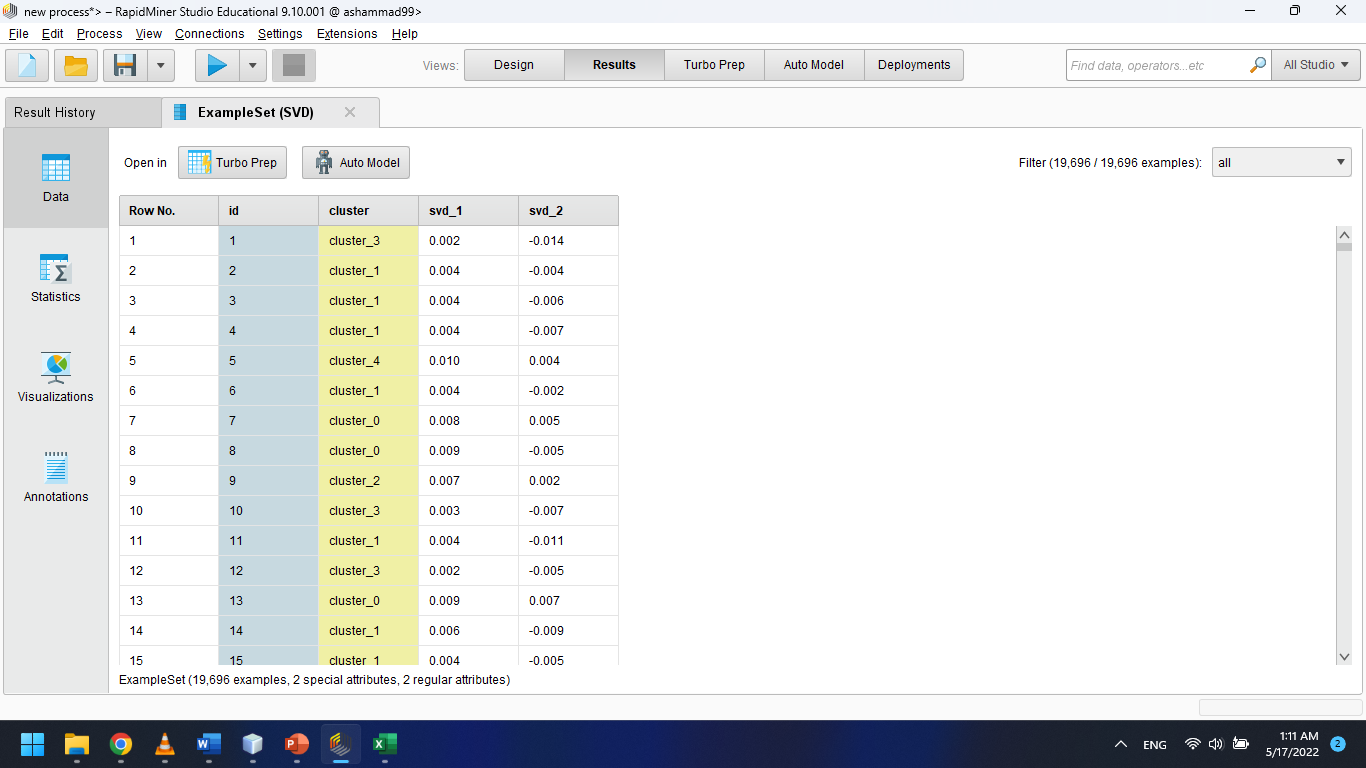
**- Select Attribute:** select 5 Attributes (age, Ease of online booking, Flight Distance, Food and drinks, Inflight WIFI services)

**-Normalize:** normalize attribute (flight distance) in range 0 and 5

**-Clustering: K-means, k = 5**

**-SVD: to merge 5 attributes into 2 attributes, dimensions=2**

**Screenshot of process**

**Screenshot of Result**

**Screenshot of Graph:**

