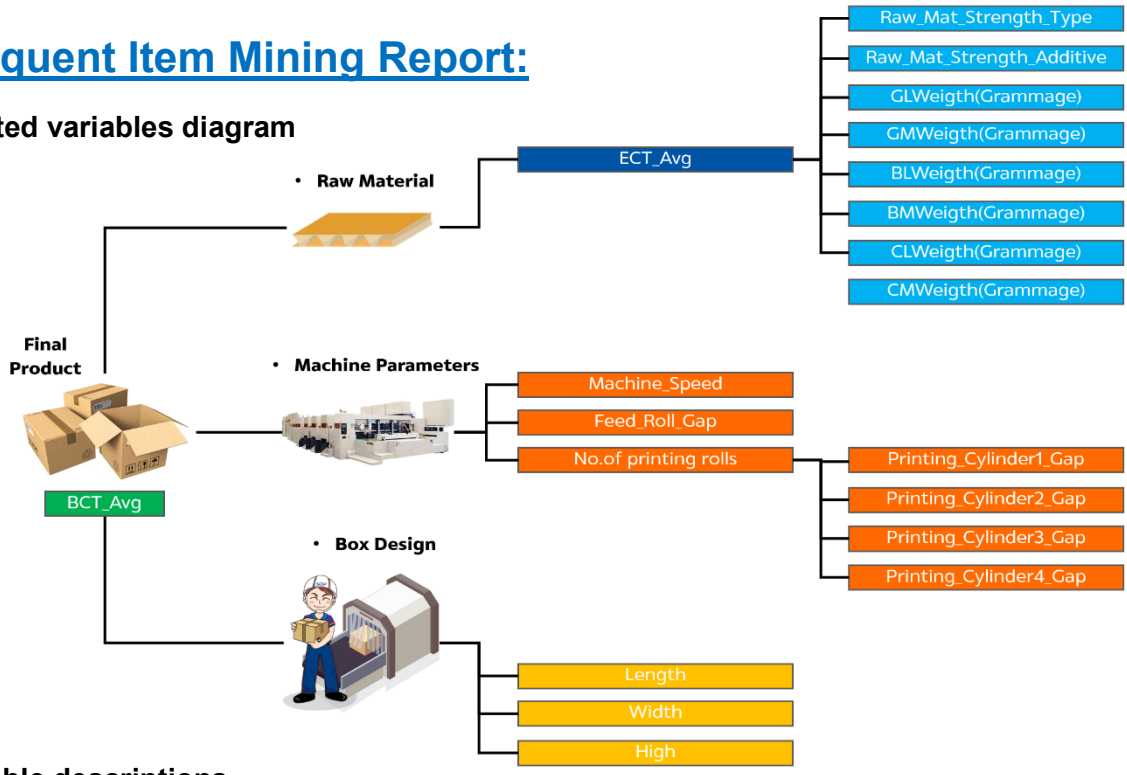
Frequent Item Mining Report:

Related variables diagram



Variable descriptions

No	Data Column	Description	Datatype	Note
1	Index	Unique ID1	Int	
2	Rwf	Unique ID2	Int	
3	Order_Number	Order Number / Lot Number	Nvarchar	
4	BCT1	The strength value of final product	Float	
5	BCT2		Float	
6	BCT3		Float	
7	BCT4		Float	
8	BCT5		Float	
9	BCT_Avg		Float	
10	BCT_Spec	Controlled parameter of the strength value of the final product. BCT from a final product does usually not exceed the BCT Spec	Float	
11	ECT1	The strength value of raw materials that affect directly to the final product	Float	
12	ECT2		Float	
13	ECT3		Float	
14	ECT4		Float	
15	ECT5		Float	
16	ECT_Avg		Float	
17	SKU	Type of product	Nvarchar	
18	Speed_Act	Production speed while produce a product	Float	
19	Feed_Roll_Gap	Width applied by a gap feeding cylinder roll on a raw material during production process, forming a final product.	Float	
20	Printing_Cylinder_Gap_1	Width applied by a printing cylinder roll on a raw material during the production process, forming a final product.	Float	
21	Printing_Cylinder_Gap_2		Float	
22	Printing_Cylinder_Gap_3		Float	
23	Printing_Cylinder_Gap_4		Float	
24	Number_of_use_Printing_Cylinder	Number of Printing Cylinder are ultilized to product a product	Int	If the number of Priting Cylinder is high, the strength value of the final product will be reduced
25	Raw_Mat_Strength_Type	B = Lower Strength C = Medium Strength BC = Super Strength	Nvarchar	
26	Raw_Mat_Strength_Additive	Level of the additive chemical that can improve corrugated box strength G0=0% G1=5% G2=10% G3=15% G4=20% G5=25% G6=30%	Nvarchar	
27	Raw_Mat_Combination	Combination of raw materail	Nvarchar	
28	GL	Layer 1 of raw materail	Nvarchar	In case of Raw_Mat_Strength_Type = B CM should be Null CMWeigth(Grammage) should be Null CL should be Null CLWeigth(Grammage) should be Null because a raw materail does not have a combination of strenght type C ---------- In case of Raw_Mat_Strength_Type = C BM should be Null BMWeigth(Grammage) should be Null BL should be Null BLWeigth(Grammage) should be Null because a raw materail does not have a combination of strenght type B
29	GLWeigth(Grammage)	Layer 1's weight of raw materail in grammage	Int	
30	BM	Layer 2 of raw materail	Nvarchar	
31	BMWeigth(Grammage)	Layer 2's weight of raw materail in grammage	Int	
32	BL	Layer 3 of raw materail	Nvarchar	
33	BLWeigth(Grammage)	Layer 3's weight of raw materail in grammage	Int	
34	CM	Layer 4 of raw materail	Nvarchar	
35	CMWeigth(Grammage)	Layer 4's weight of raw materail in grammage	Int	
36	CL	Layer 5 of raw materail	Nvarchar	
37	CLWeigth(Grammage)	Layer 5's weight of raw materail in grammage	Int	
38	Finished_Product_Wid	Width of the final product	Float	The strength value of the final product also related to the product's height. A product that more height is lower strength than a shorter one
39	Finished_Product_Leg	Length of the final product	Float	
40	Finished_Product_Hig	Higth of the final product	Float	

Create Binary and Dummy Column:

We can apply a condition to create binary variable columns, showing you the steps below. Moreover, some category columns will bel convert categorical variables into binary variables. To do this, we will use the dummy method, which involves creating a new column for each unique category in the categorical variable. Each column will have a value of 0 or 1, indicating whether the category is present or not.

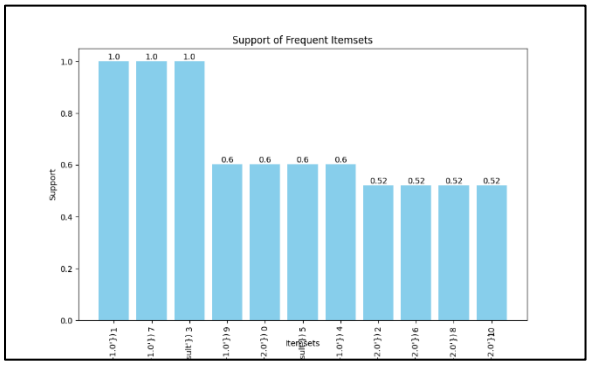
1) Generated 4 binary columns with conditions
    a. Rejection_Result: BCT_Avg < BCT_Spec or %Diff between BCT_Avg and BCT_Spec < 3% = 0 else 1
    b. Finished_Product_Hig : Box with high between 300-600 = 1 else = 0
    c. Machine_Speed: < 350
    d. ECT_Avg: > 1.98

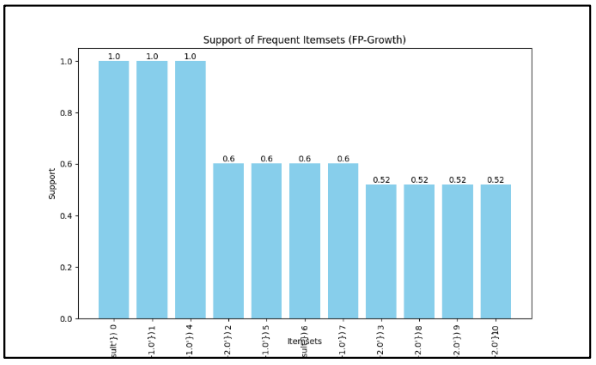
Finished_Product_Hig	Machine_Speed	ECT_Avg	Rejection_Result
0	1	1	1
0	1	1	0
0	1	1	0
0	1	1	0
0	1	1	0
0	1	1	0

2) Created binary columns by using the dummy method, getting 30 binary columns.

Raw_Mat_Strength_Type_B	Raw_Mat_Strength_Type_BC	Raw_Mat_Strength_Type_C	Raw_Mat_Strength_Additive_G0	Raw_Mat_Strength_Additive_G1	Raw_Mat_Strength_Additive_G2
1	0	0	1	0	0
0	0	1	1	0	0
1	0	0	0	0	0
0	0	1	0	0	0
0	0	1	0	0	0
0	0	1	0	0	0
1	0	0	1	0	0
0	1	0	0	0	0
0	0	1	0	0	0

3) Seleted data with Rejection_Result = 1 and applied Apriori and FP-Growth function to the new dataset with binary columns (vary min_support from 0.3-0.5 for filter data)

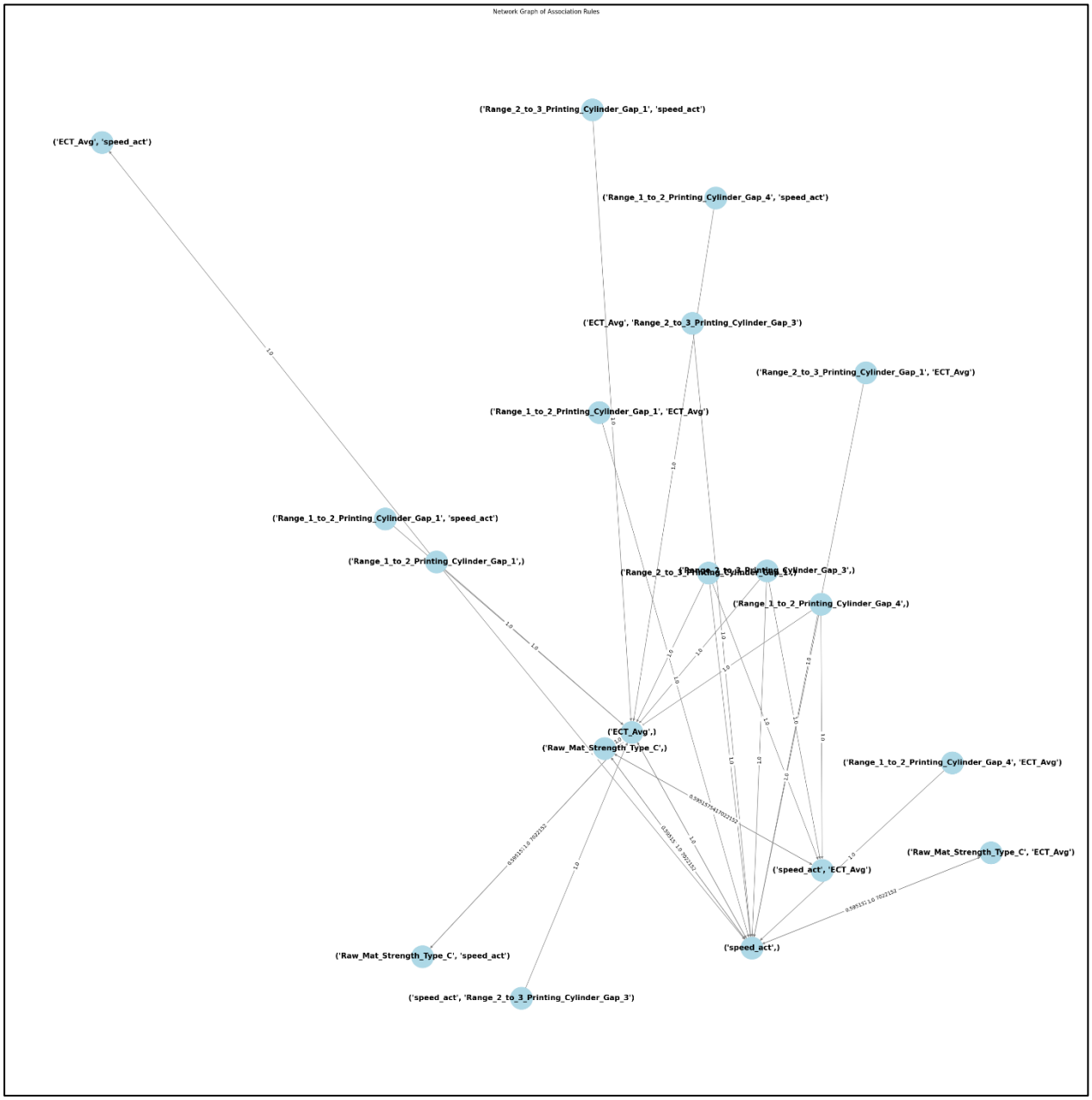




Conclusion:

The interesting patterns that we found can be interpreted that to ensure that the box strength meets the specification, the fp-growth and apriori data mining pattern methods recommend 5 major variables from 40 variables by using Network Graph of Association Rules:

1) Raw material strength (ECT_Avg).
2) Printing cylinder gap (Printing_Cylinder_Gap 1-4).
3) Feed roll gap (Feed_Roll_Gap).
4) Raw material type (Raw_Mat_Strength_Type).
5) Additive type (Raw_Mat_Strength_Additive_Type).



What's Interesting We Found:

For example, from the dataset of products produced using three printing cylinders, the FP-growth and Apriori algorithms suggest the following:

1) Controlling the printing cylinder gap of each printing cylinder unit within 0.0-1.0 and the average of all printing units should be between: 1.0-2.0.
2) Controlling the feed roll gap between: 1.0-2.0
3) Raw material type: C.
4) Incorporating additive type: G3.
5) Raw material strength: > 1.98

To verify the suggested result that the five variables affect the final product according to box production fundamentals (ISO Standard), we will select the sample data from product SKU "1397-123-00G3" (the top-selling product) following the five specified data pattern conditions suggested by the FP-growth and Apriori algorithms. We will then use a scatter plot to analyze the trend of BCT_Avg. It can be inferred that the pattern data of the printing cylinder and feed roll gaps applied to the raw material surface, raw material type, additive type, and raw material strength all have an impact on the final product strength (BCT_Avg), as observed in the actual production results.

Specifically, when the gap between the printing cylinders is low (indicating high pressure), the BCT_Avg is low. Conversely, when the gap is high (indicating low pressure), the BCT_Avg is high. This suggests that controlling the printing cylinder gaps can help optimize the product strength. Additionally, the raw material type, additive type, and raw material strength also influence the BCT_Avg, and these variables should be considered when developing the production recipe.

By applying these condition patterns and analyzing the scatter plot, we can gain insights into how different combinations of variables affect the final product strength.