DATA PREPROCESCING

Cleaning the 'Education'

Education	Children	Pet	Absenteeism Time in Hours
1	2	1	4
1	1	0	0
1	0	0	2
1	2	0	4
1	2	1	2
1	0	0	2
1	1	4	8
1	4	0	4
1	2	0	40
3	1	1	8

Education	Children	Pet	Absenteeism Time in Hours
0	2	1	4
0	1	0	0
0	0	0	2
0	2	0	4
0	2	1	2
0	0	0	2
0	1	4	8
0	4	0	4
0	2	0	40
1	1	1	8

1	High school	538
2	Graduate	73
3	Postgraduate	40
4	A master or a doctor	4

1. map 'Education' variables; the result is a dummy
df['Education'] = df['Education'].map({1:0, 2:1, 3:1, 4:1})

0	High school	538
1	Not High School	117

ANALYZING REASON VS PROBABILITY - TABLEAU