

## DATA MINING CUP 2019

### Features

Column name	Description	Value range
<b>trustLevel</b>	A customer's individual trust level. 6: Highest trustworthiness	{1,2,3,4,5,6}
<b>totalScanTimeInSeconds</b>	Total time in seconds between the first and last product scanned	Positive whole number
<b>grandTotal</b>	Grand total of products scanned	Positive decimal number with maximum two decimal places
<b>lineItemVoids</b>	Number of voided scans	Positive whole number
<b>scansWithoutRegistration</b>	Number of attempts to activate the scanner without actually scanning anything	Positive whole number or 0
<b>quantityModification</b>	Number of modified quantities for one of the scanned products	Positive whole number or 0
<b>scannedLineItemsPerSecond</b>	Average number of scanned products per second	Positive decimal number
<b>valuePerSecond</b>	Average total value of scanned products per second	Positive decimal number
<b>lineItemVoidsPerPosition</b>	Average number of item voids per total number of all scanned and not cancelled products	Positive decimal number
<b>fraud</b>	Classification as fraud (1) or not fraud (0)	{0,1}