**Zero - Variance Features**

Instruction

Please ensure you update all the details:

**Name: RAKESH SETHU NP**

**Batch Id: 11052022\_7.30PM**

**Topic: Data Pre-Processing**

Variance measures how far a set of data is spread out. A variance of zero indicates that all the data values are identical. There are various techniques to remove this for transforming the data into the suitable one for prediction.

**Problem statement:**

Find which columns of the given dataset with zero variance, explore various techniques used to remove the zero variance from the dataset to perform certain analysis.



|  |  |  |  |
| --- | --- | --- | --- |
| Name of Feature | Description | Type | Relevance |
| id | id | Nominal | Not a useful Information |
| Square.length | Length of square | Continuous, Ratio | Useful Information |
| Square.breadth | Breadth of square | Continuous, Ratio | Useful Information |
| Rec.length | Length of rectangle | Continuous, Ratio | Useful Information |
| Rec.breadth | Breadth of rectangle | Continuous, Ratio | Useful Information |
| colour | color | Nominal | Not a useful Information |

A picture containing shape, arrow

Description automatically generated

**Hints:**

For each assignment, the solution should be submitted in the below format

1. Work on each feature of the dataset to create a data dictionary as displayed in the below image:



1. Consider the Z\_dataset.csv dataset
2. Research and perform all possible steps for obtaining solution
3. All the codes (executable programs) should execute without errors
4. Code modularization should be followed
5. Each line of code should have comments explaining the logic and why you are using that function