Differentiate genuine and counterfeit notes

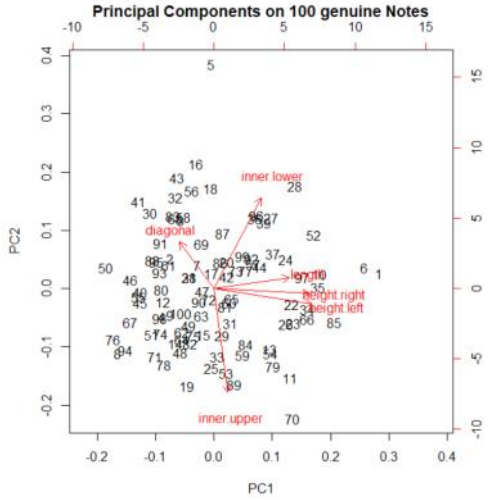
Monday, March 30, 2020 5:53 PM

Analysis on Swiss bank notes to classify genuine and counterfeit notes

Observations :

- The provided dataset makes quantitative measures of Swiss bank notes.
- There are 100 genuine and 100 counterfeit notes
- The six variables are length of the bank note, height of the bank note, measured on the left, height of the bank note measured on the right, distance of the inner frame to the lower border, distance of inner frame to upper border, and length of the diagonal.

Principal Component Analysis on Genuine Notes

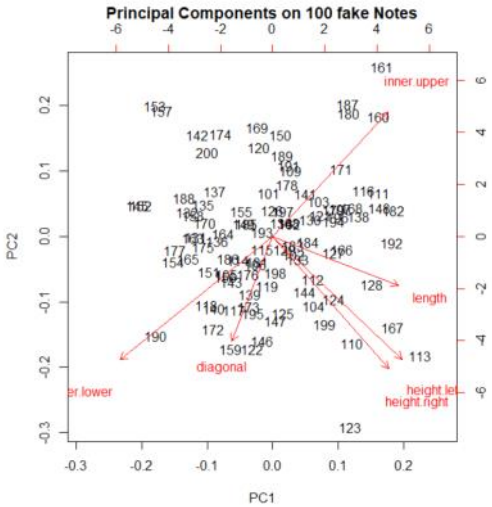


	PC1	PC2
length	0.44955352	0.07399926
height.left	0.58504531	-0.10733153
height.right	0.57218242	-0.03575517
inner.lower	0.28120306	0.61627568
inner.upper	0.08238938	-0.70885999
diagonal	-0.20583422	0.31535285

Inference :

i. Height.left, Height.right and length are the major contributors to PC1 hence we can say that they vary the most
ii. Inner.upper and inner.lower are unrelated to height.left and height.right as the vectors are almost orthogonal in the Biplot.
iii. Inner.upper and inner.lower are major contributors to Principal Component 2
iv. We can see a cluster being formed with all the Notes in close vicinity and no outliers

Principal Component Analysis on Fake Notes

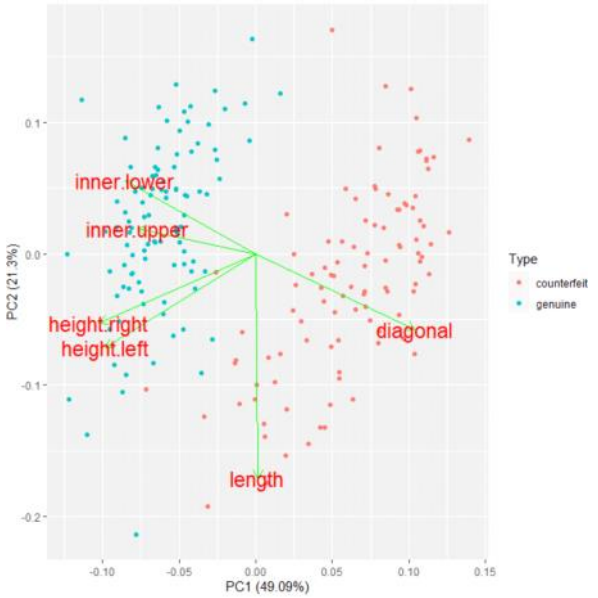


	PC1	PC2
length	0.4340833	-0.1775842
height.left	0.4456165	-0.4433000
height.right	0.3989756	-0.4785602
inner.lower	-0.5256536	-0.4452589
inner.upper	0.3972111	0.4491523
diagonal	-0.1404554	-0.3780684

Inference

i. The variable which shows maximum variation in Fake Notes is inner.lower. It contributes the most to principal Component 1 and is also one of the major contrinutors to principal component 2
ii. The major contributor to Principal component 2 is height.right
iii. We see many outliers in the biplot for example, Note 161 has an unusual large inner.upper score and Note 123 has a large height.right

Principal Component Analysis on Genuine and Counterfeit notes



	PC1	PC2
length	0.006987029	-0.81549497
height.left	-0.467758161	-0.34196711
height.right	-0.486678705	-0.25245860
inner.lower	-0.406758327	0.26622878
inner.upper	-0.367891118	0.09148667
diagonal	0.493458317	-0.27394074

Inference :

i. PC1 accounts for 66.75% of variation in the dataset and PC2 accounts for 20.82% in the dataset. Together PC1 and PC2 account for 88% variation in the dataset.
ii. On plotting the fake and genuine notes in PC1 and PC2 axis, we see a clear separation between the two kinds of notes.
iii. Diagonal is the most differentiating factor which differentiates genuine and counterfeit notes. Counterfeit notes are on the positive side of PC1 axis for diagonal and Genuine notes are on the negative side.
iv. It doesn't come as a surprise that Diagonal contributes the most to PC1. The other contributors to PC1 are height.right and height.left