

Ankit Nigam

GCDAl – August batch 2019

# Gender recognition by voice

# Dataset information

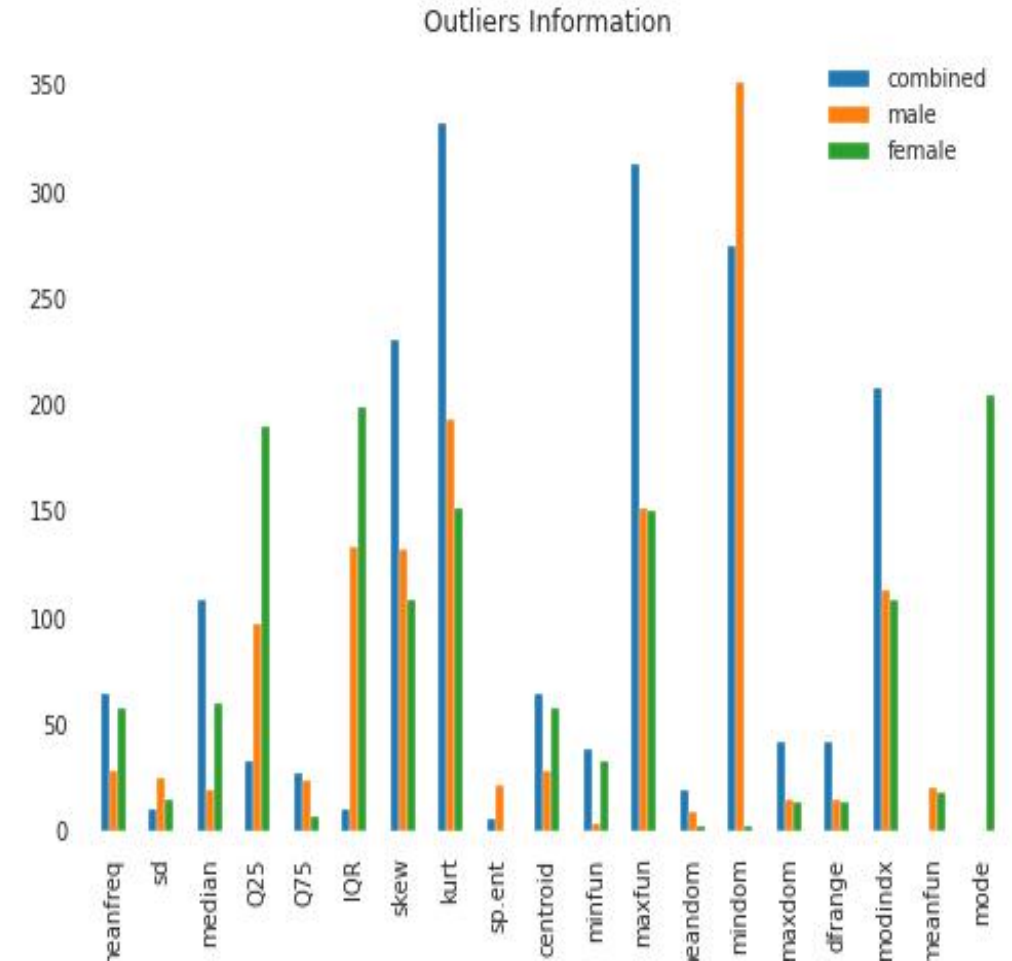
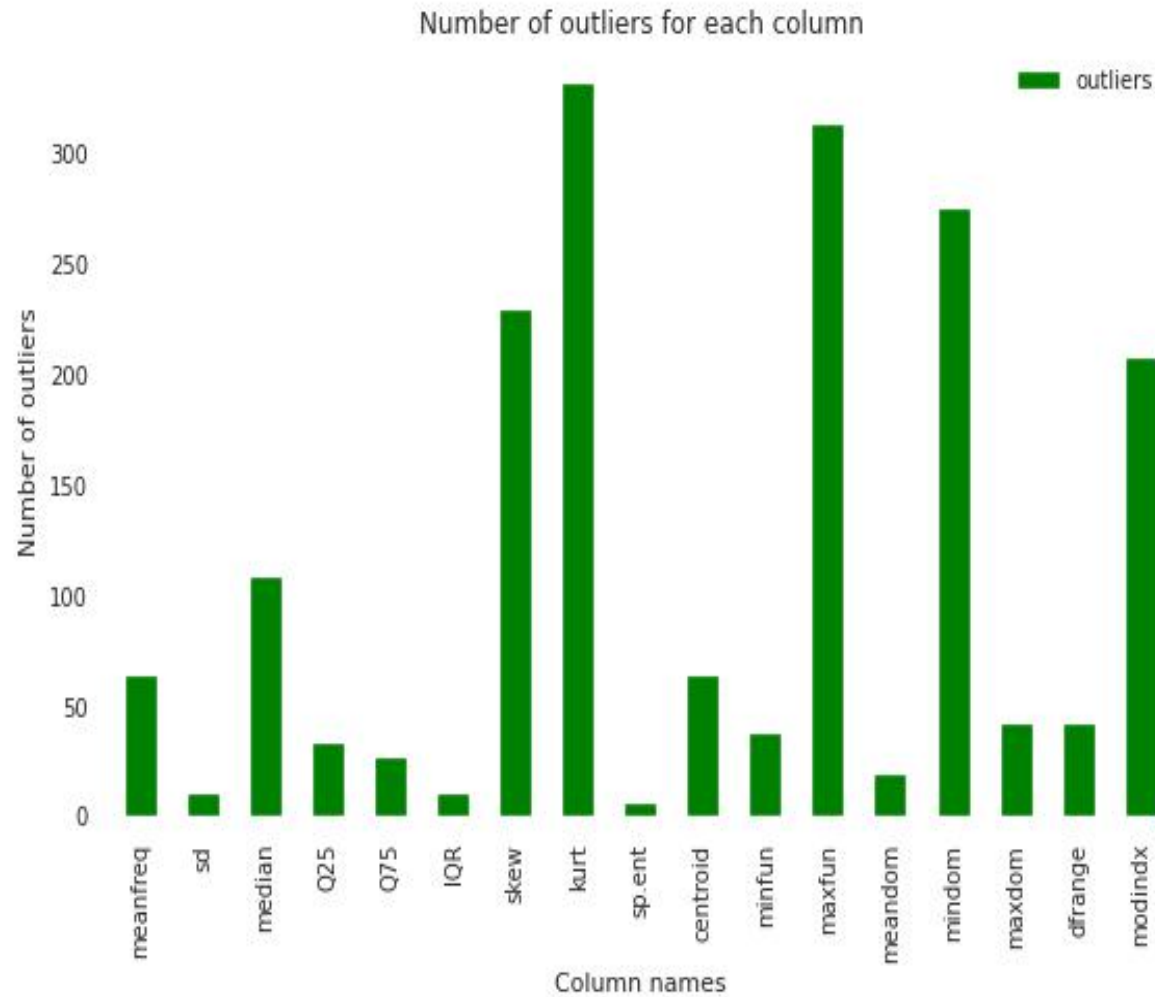
## Input features

Fields	Purpose
meanfreq	mean frequency (in kHz)
sd	standard deviation of frequency
median	median frequency (in kHz)
Q25	first quantile (in kHz)
Q75	third quantile (in kHz)
IQR	interquantile range (in kHz)
skew	skewness
kurt	kurtosis
sp.ent	spectral entropy
sfm	spectral flatness
mode	mode frequency
centroid	frequency centroid
peakf	peak frequency
meanfun	average of fundamental frequency measured across acoustic signal
minfun	minimum fundamental frequency measured across acoustic signal
maxfun	maximum fundamental frequency measured across acoustic signal
meandom	average of dominant frequency measured across acoustic signal
mindom	minimum of dominant frequency measured across acoustic signal
maxdom	maximum of dominant frequency measured across acoustic signal
dfrange	range of dominant frequency measured across acoustic signal
modindx	modulation index

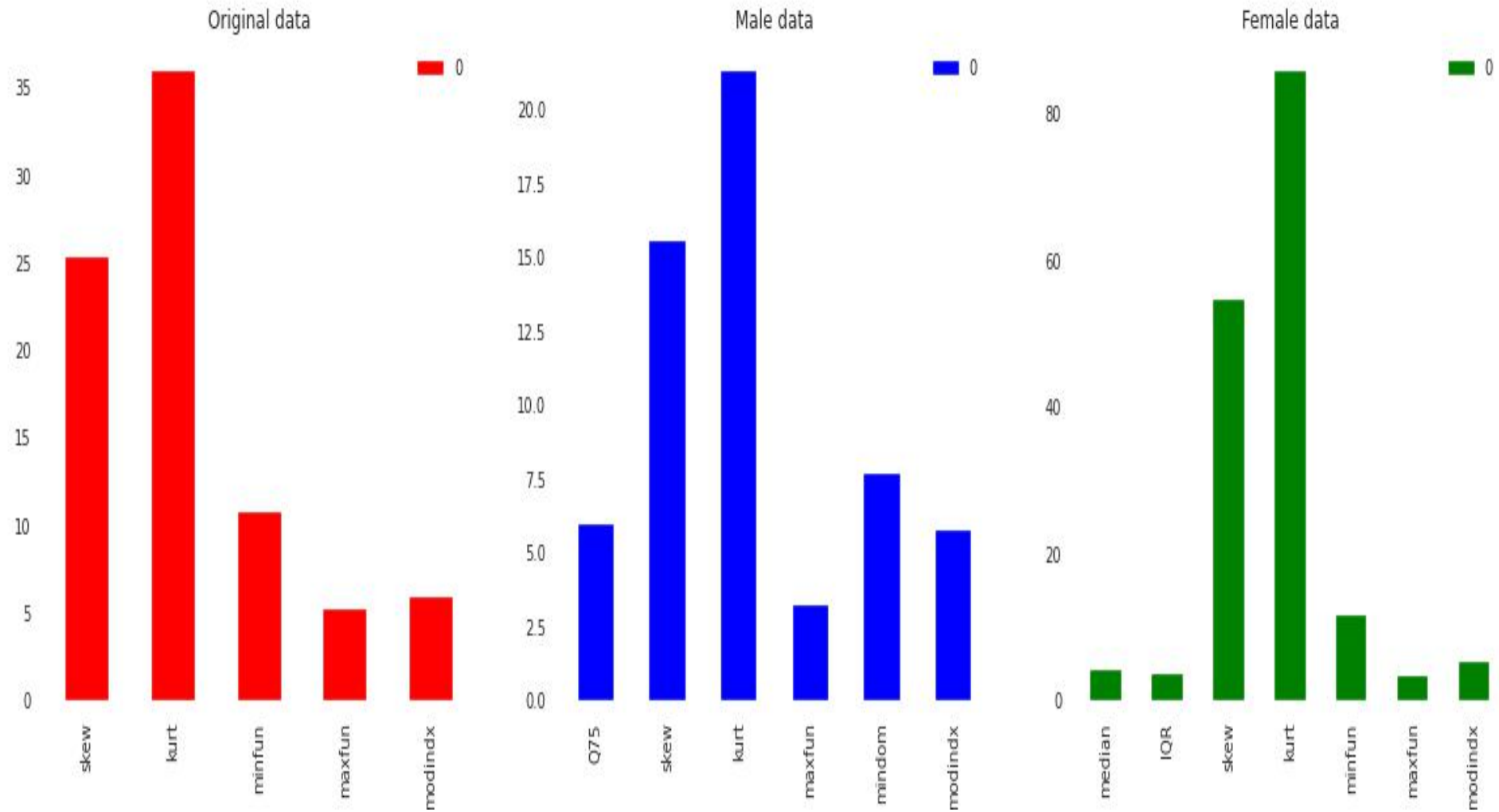
## Target

Fields	Purpose
label	male or female

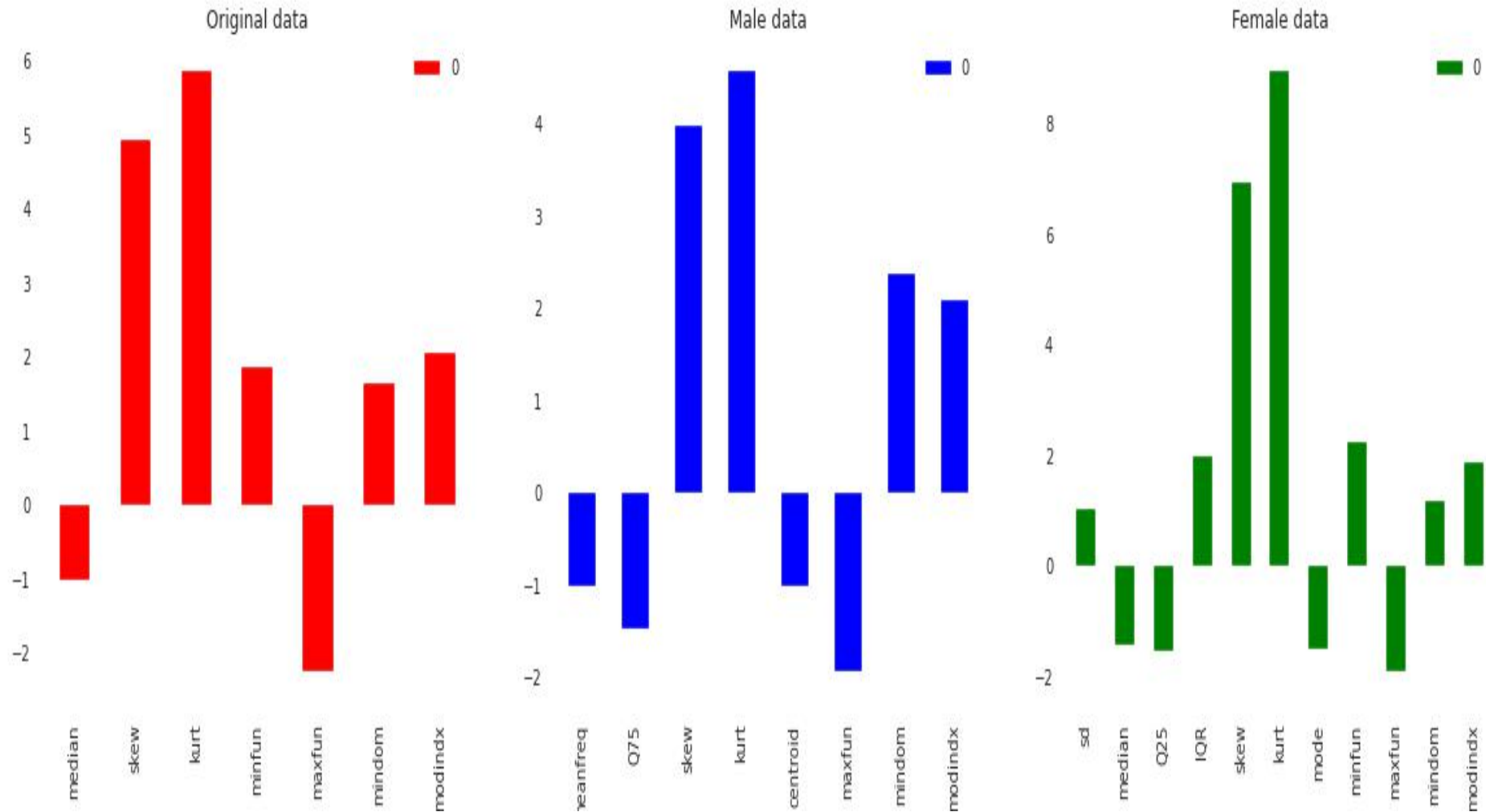
# Outlier Information



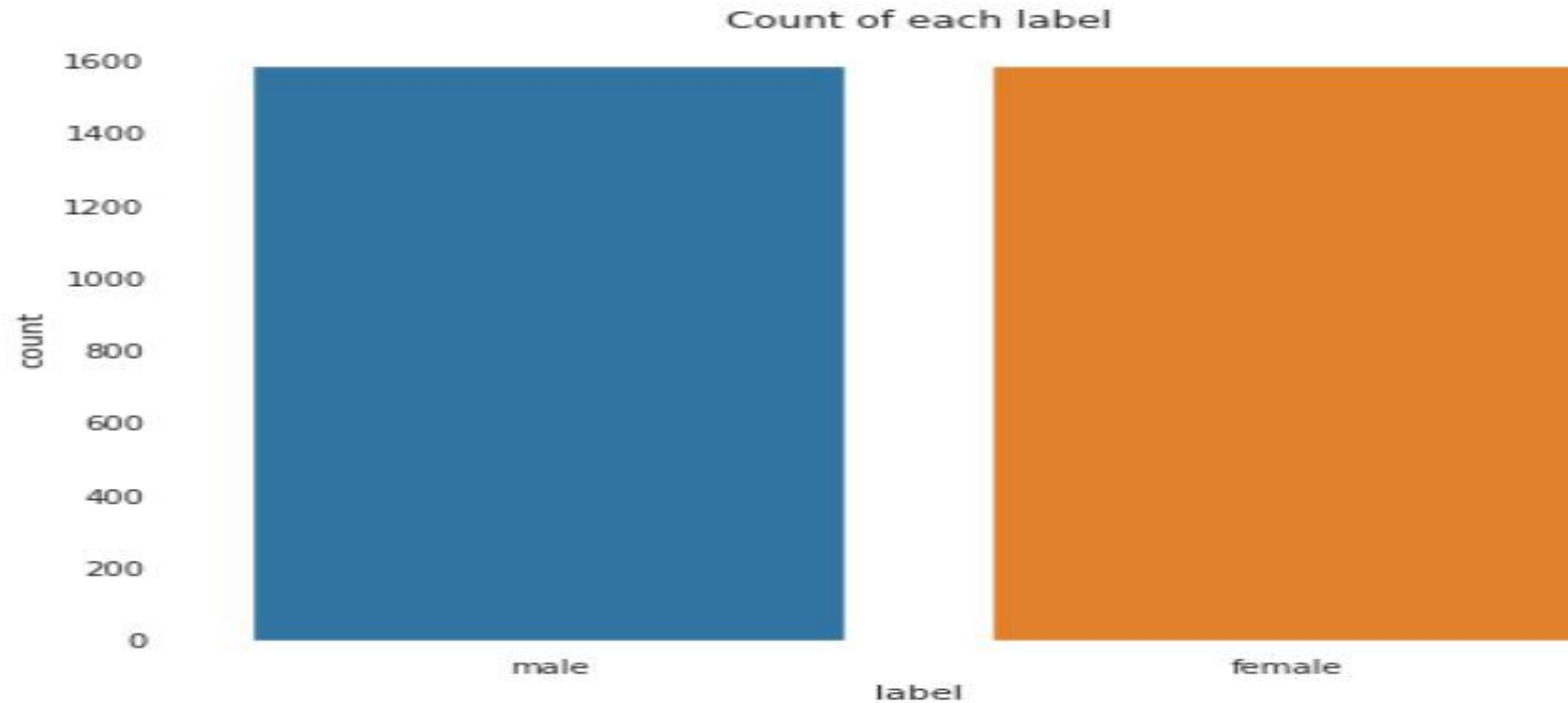
# Data distribution – Kurtosis



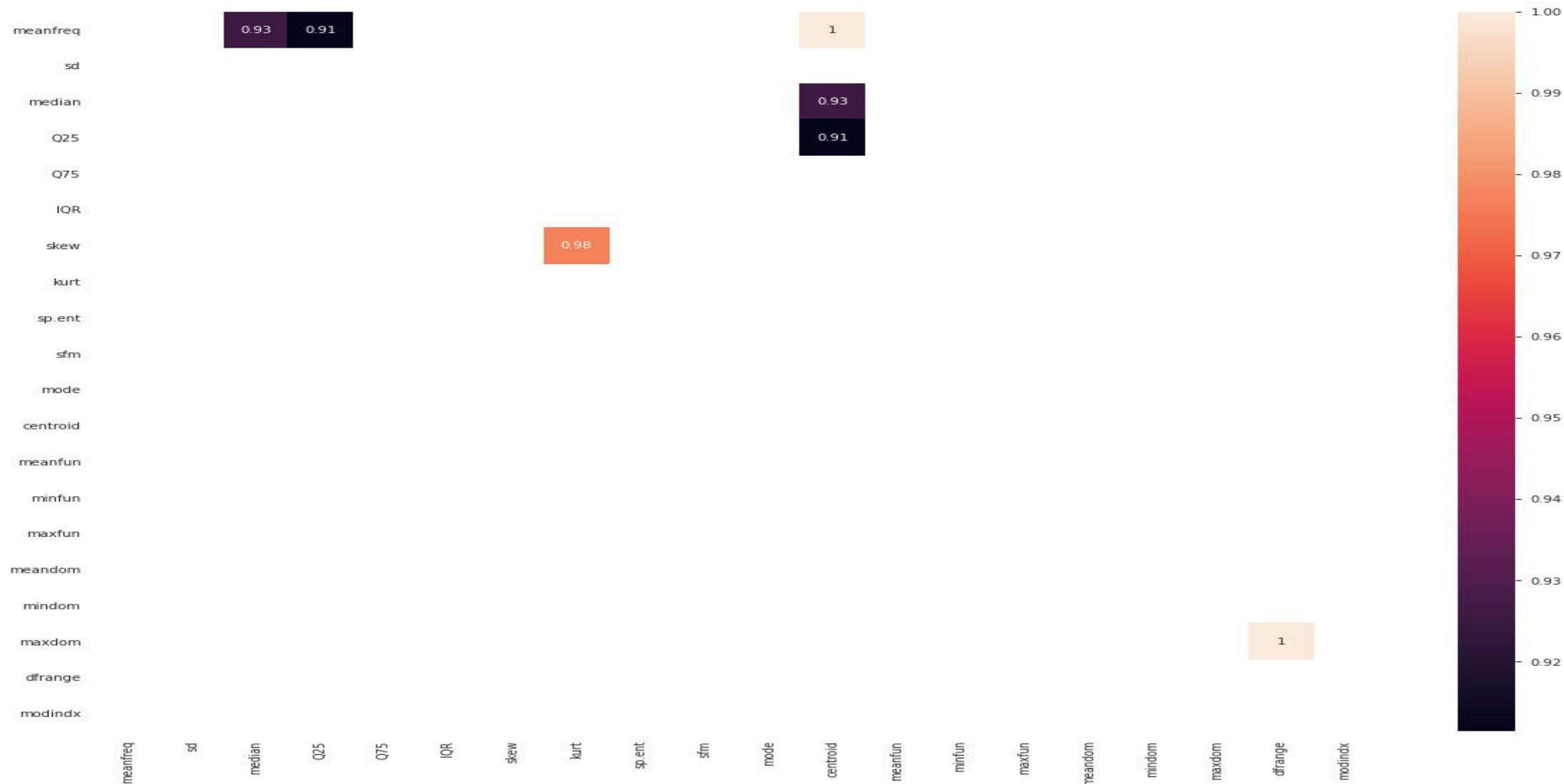
# Data distribution – Skewness



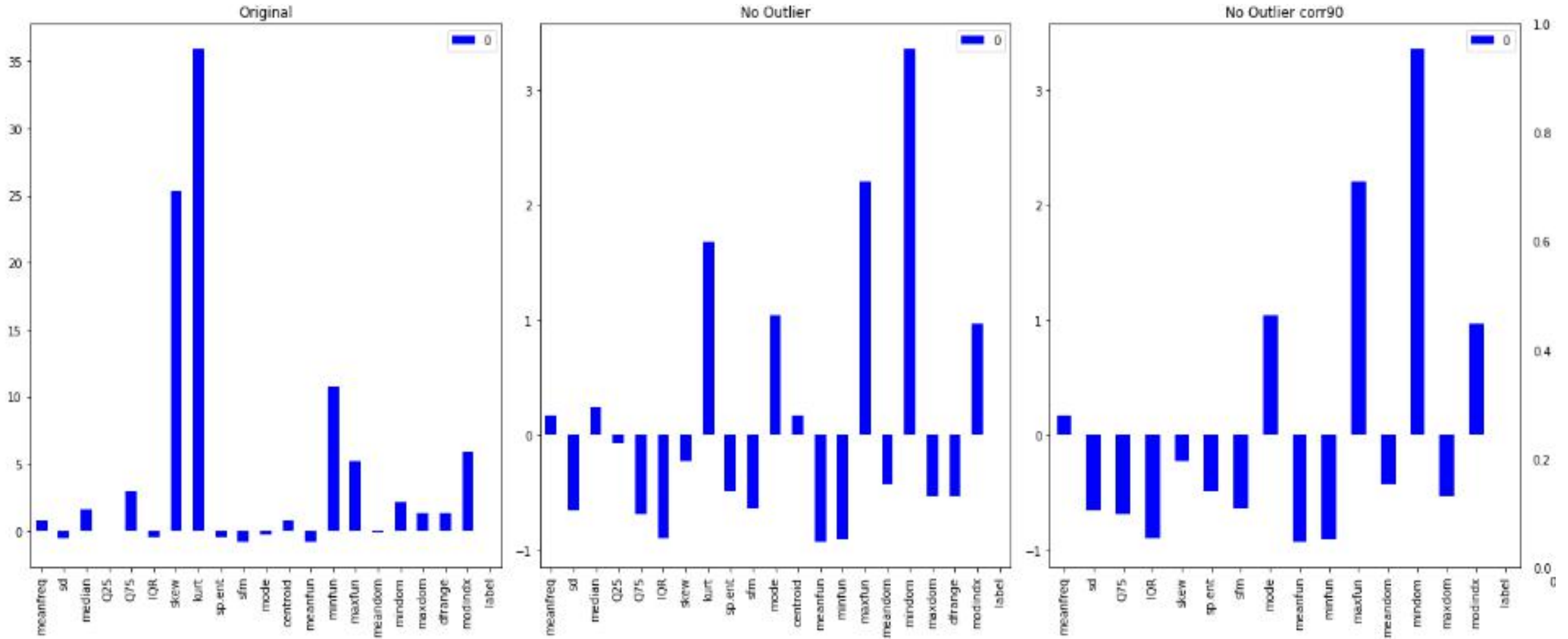
# Data distribution – target



# Correlation between columns

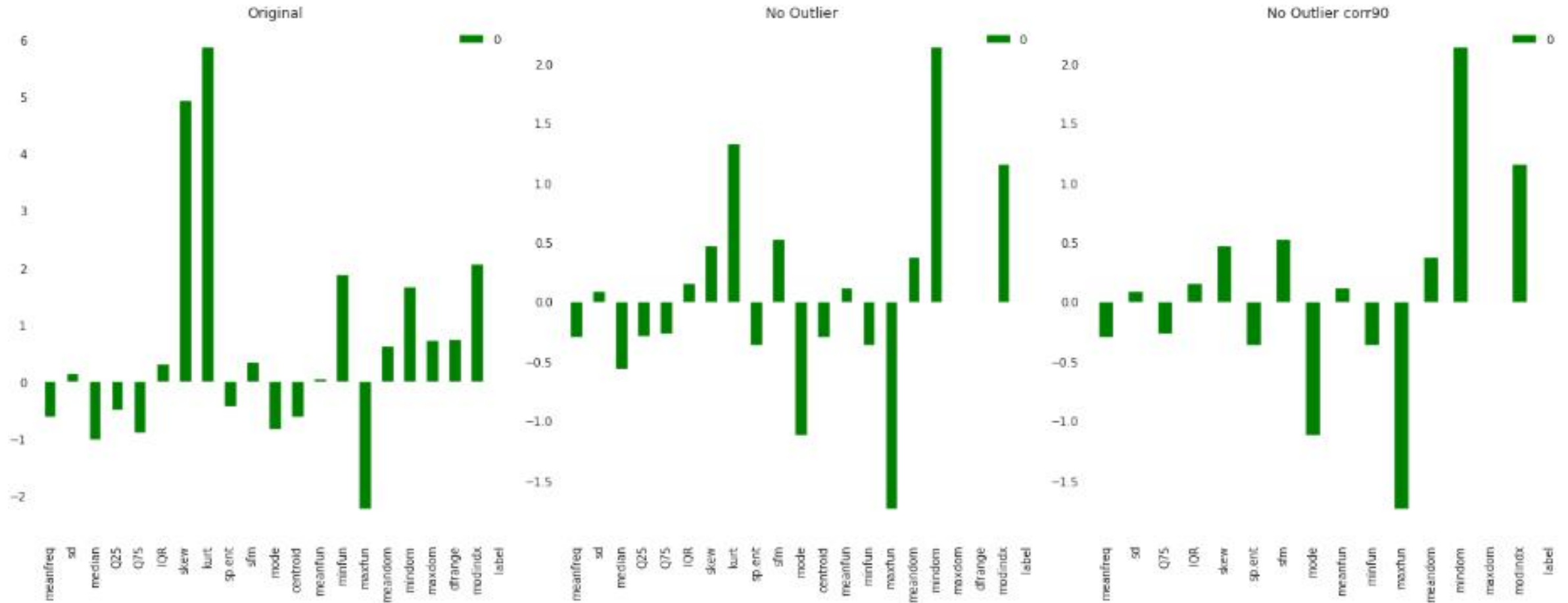


# Data distribution no outliers – Kurtosis

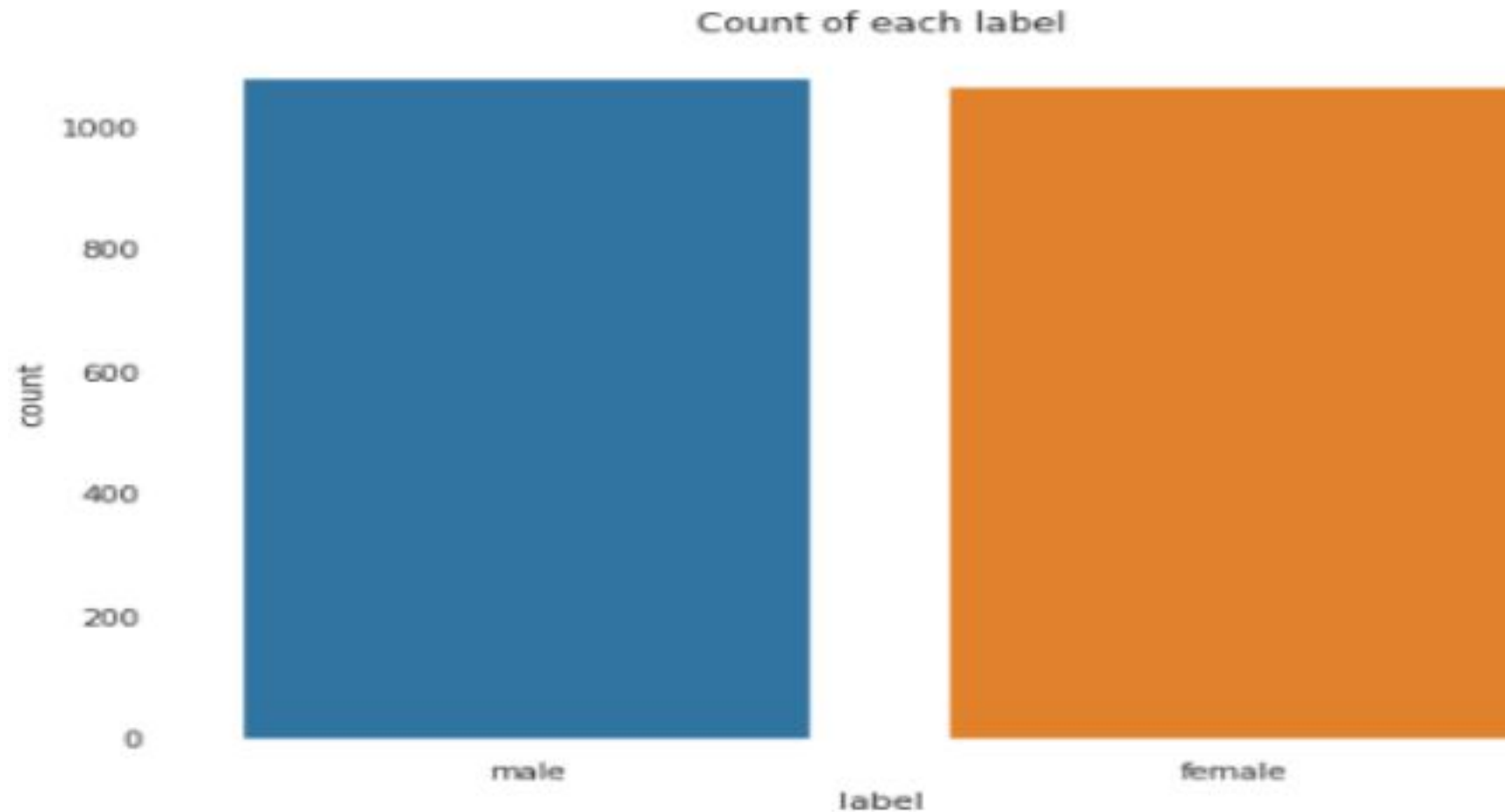




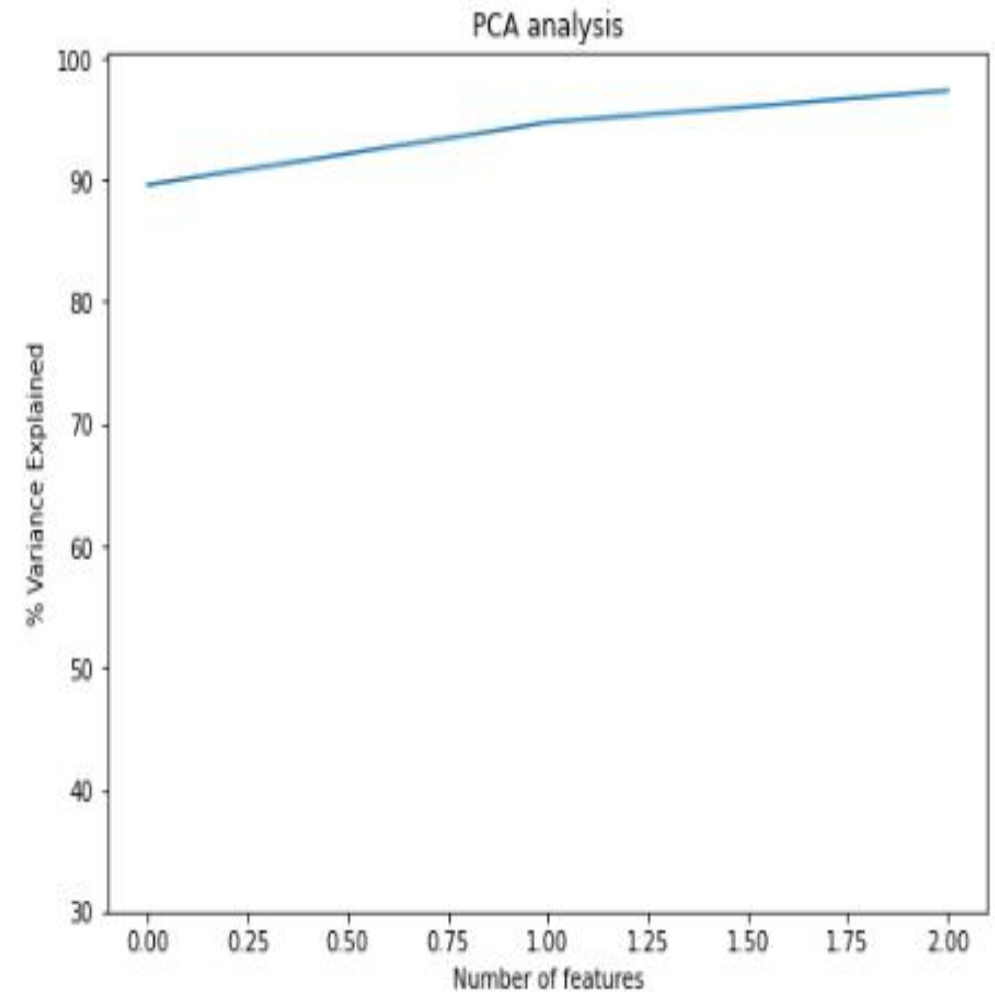
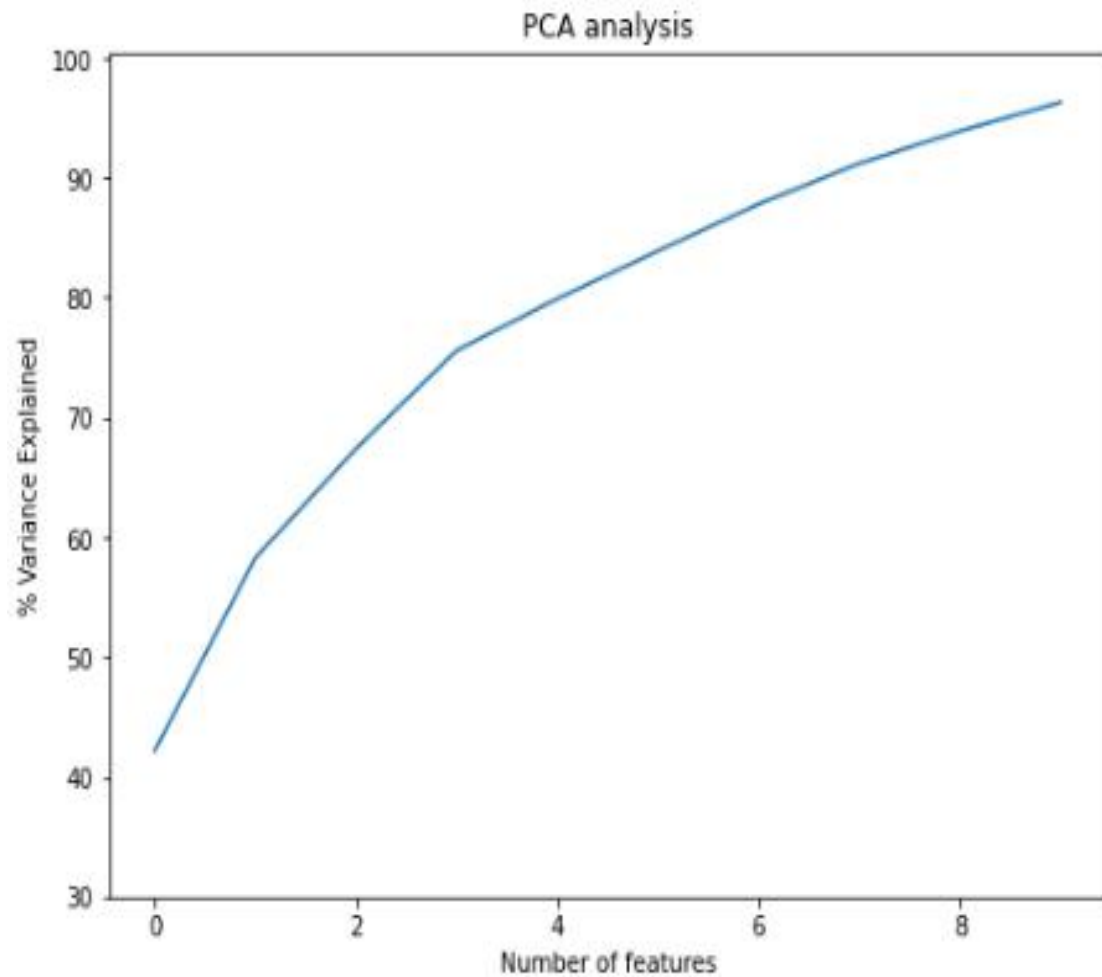
# Data distribution no outliers – Skewness



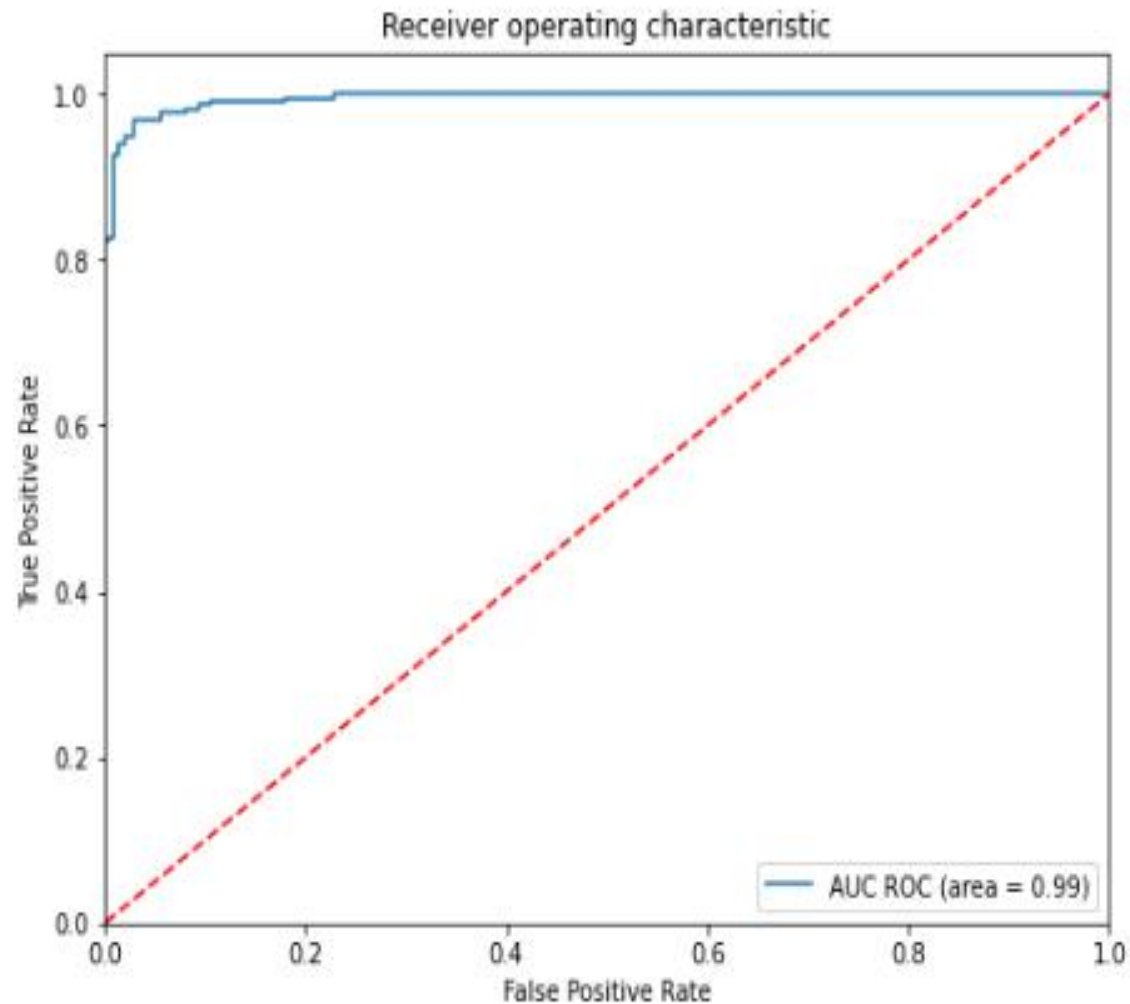
# Data distribution after outlier removal – target



# Dimensionality reduction - PCA



# Model Selection using pipeline



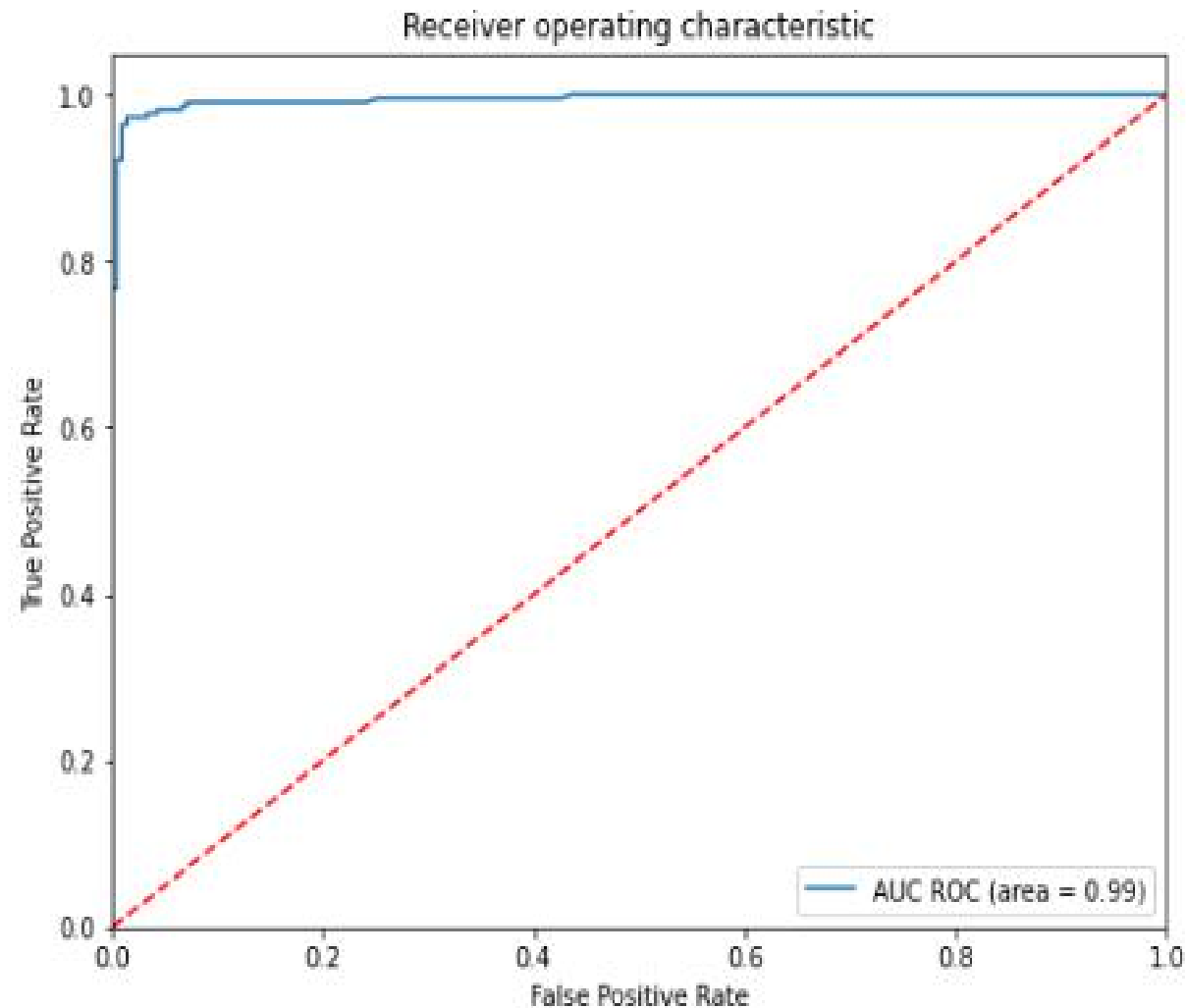
Confusion Matrix:

	male	female
male	206	5
female	11	208

Classification Report :

	precision	recall	f1-score	support
0	0.95	0.98	0.96	211
1	0.98	0.95	0.96	219
accuracy			0.96	430
macro avg	0.96	0.96	0.96	430
weighted avg	0.96	0.96	0.96	430

# Model Selection using pipeline



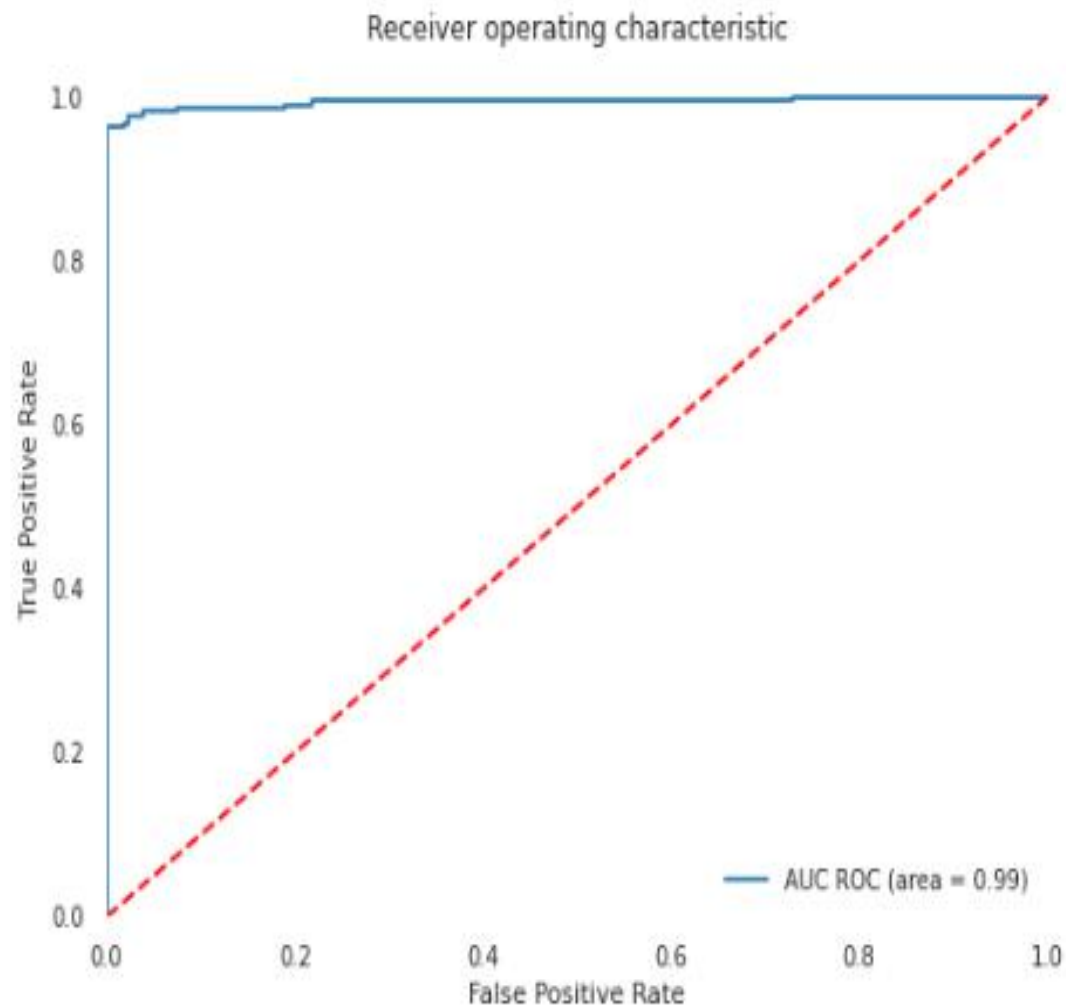
Confusion Matrix:

	male	female
male	209	2
female	9	210

Classification Report :

	precision	recall	f1-score	support
0	0.96	0.99	0.97	211
1	0.99	0.96	0.97	219
accuracy			0.97	430
macro avg	0.97	0.97	0.97	430
weighted avg	0.97	0.97	0.97	430

# Final Model – Support Vector Classifier



Confusion Matrix:

	male	female
male	209	2
female	8	211

Classification Report :

	precision	recall	f1-score	support
0	0.96	0.99	0.98	211
1	0.99	0.96	0.98	219
accuracy			0.98	430
macro avg	0.98	0.98	0.98	430
weighted avg	0.98	0.98	0.98	430

Thank You