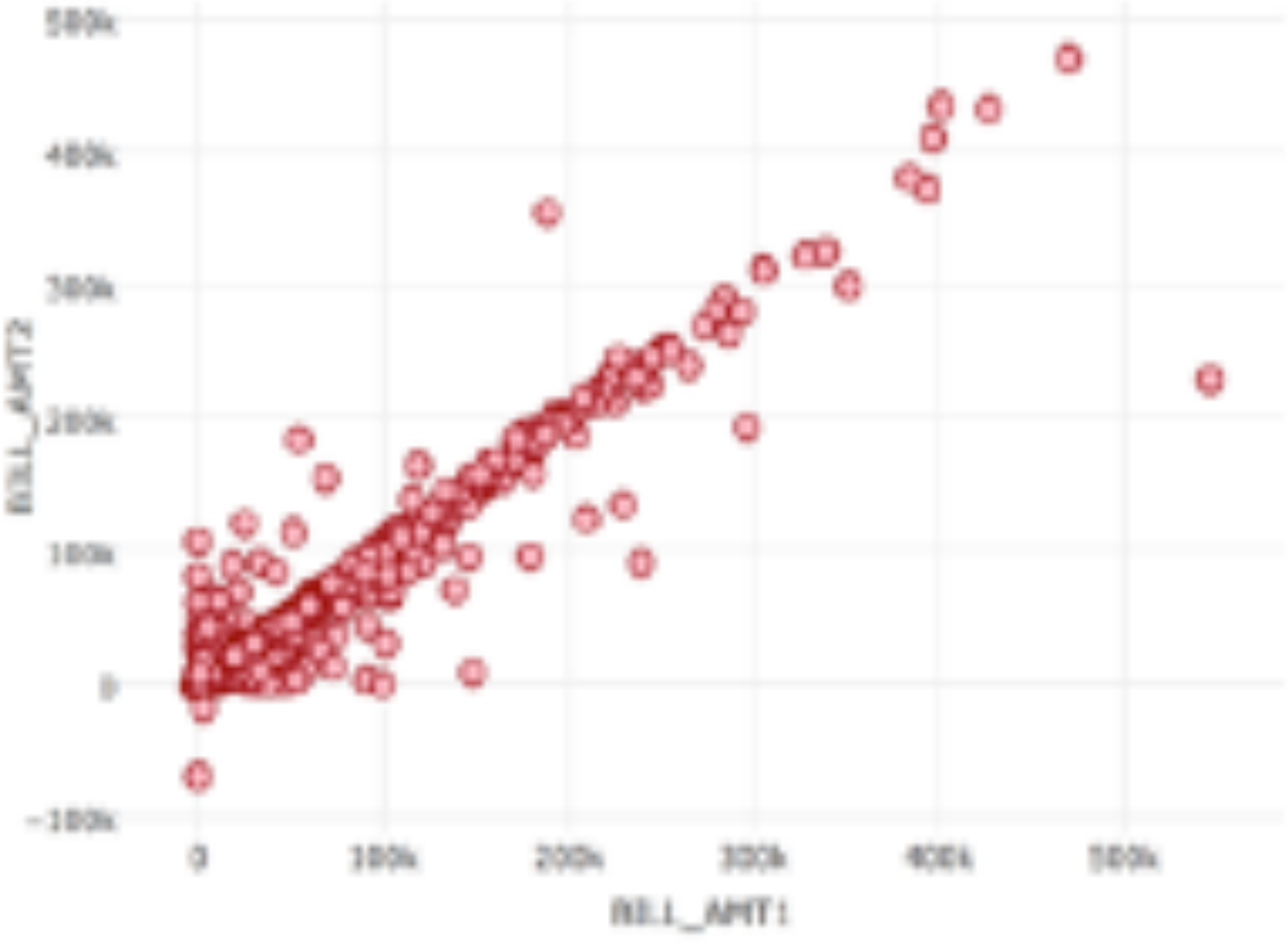
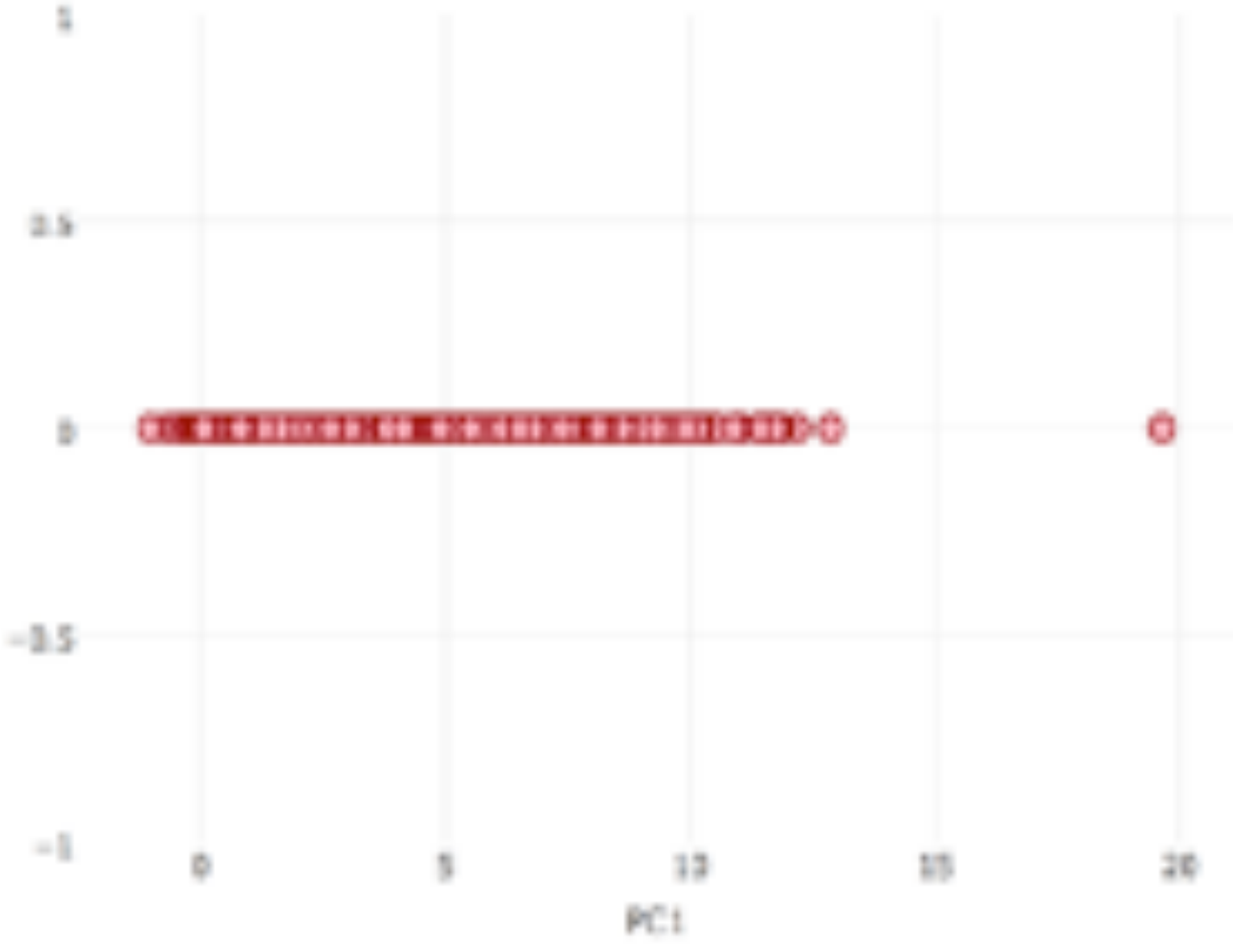

H2O.ai

Principal Components Analysis

- Orthogonal rotation of covariance or correlation matrix that orders derived measures from highest to lowest variation
- Useful for dimensionality reduction / removing collinearities


500k
400k
300k
200k
100k
0
-100k
BILL_AMT2
0
100k
200k
300k
400k
500k
BILL_AMT1


1
0.5
0
−0.5
−1
0
5
10
15
20
PC1

Pros and Cons of Principal Components Analysis

Pros

- Can be computed using covariance/correlation matrix

Cons

- Sensitive to skewed measures
- Sensitive to outliers
- Categorical data requires preprocessing
  - Multiple Correspondence Analysis
  - Multi-Dimensional Scaling

Principal Components Analysis

- Orthogonal rotation of covariance or correlation matrix that orders derived measures from highest to lowest variation
- Useful for dimensionality reduction / removing collinearities