

ANÁLISE DE COMPONENTES PRINCIPAIS (PCA)

As informações a seguir apresentam uma análise utilizando a Análise de Componentes Principais (PCA) para detalhar as características das ocorrências de polícia na Baixada Santista. O PCA simplifica e interpreta eficientemente dados multidimensionais, extraindo componentes principais que capturam a variação nas medidas geográficas. Isso proporciona uma visão clara das características distintivas entre os locais das ocorrências, identificando padrões e variações.

Dimensão dos Dados

Número de instâncias: Número total de ocorrências após remoção de dados ausentes.

Número de atributos: 2 atributos numéricos preditivos: latitude e longitude.

Matriz de Covariância

A matriz de covariância é uma métrica estatística que caracteriza a interdependência entre as variáveis de um conjunto de dados, revelando as correlações entre suas características. A seguir está a Matriz Quadrada Reduzida da Matriz de Covariância:

```
[[1.00008163 1.00007313]
 [1.00007313 1.00008163]]
```

Autovalores e Autovetores

Os autovalores e autovetores, derivados da matriz de covariância, oferecem insights sobre as direções primárias dos dados e suas relevâncias. Os autovetores delineiam as direções principais, ou componentes principais, enquanto os autovalores indicam a variância explicada por cada componente.

Autovalores: [2.00015475e+00 8.49825139e-06]

Autovetores: [[0.70710678 0.70710678]
 [0.70710678 -0.70710678]]

Resultado

A plotagem do PCA das ocorrências de polícia na Baixada Santista revelou que a maioria dos incidentes tem características geográficas semelhantes, concentradas próximas à origem do gráfico. No entanto, alguns incidentes, como um ponto de "Furto de Veículo", estão dispersos, indicando ocorrências geograficamente distintas. Essa visualização facilita a identificação de padrões e áreas específicas de interesse para a segurança pública, ajudando no planejamento estratégico.