**Human Trafficking Data**

1000 Case Files

**Binary Features:**

|  |  |  |
| --- | --- | --- |
| Feature | 0 => No | 1 => Yes |
| Trafficker |  |  |
| Member of known gang |  |  |
| Gender |  |  |
| Married |  |  |
| Employed |  |  |
| Arrested |  |  |
| Personal |  |  |
| Property |  |  |
| Inchoate |  |  |
| Statutory |  |  |
| Misdemeanor |  |  |
| Felony |  |  |

**Non-binary Features:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Feature | <25 | <50 | <75 | <100 |
| Age |  |  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Feature | 0 (< high school) | 1 (high school) | 2 (college) | 3 (grad school) |
| Education |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Feature | Atlanta | Chicago | Dallas | Detroit | Las Vegas | San Diego | San Francisco | St. Louis | Tampa | DC |
| Location |  |  |  |  |  |  |  |  |  |  |

**Decision trees are models that split the data subjects into non-overlapping subsets.**

**Which features will give the best split?**