Example: USA Arrests (ISLR)

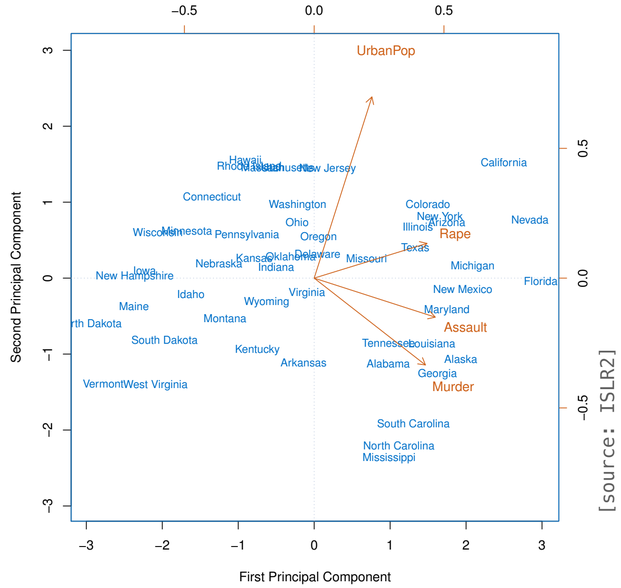
- For each state in the US:
  - number of arrests per 100 000 residents for `Assault`, `Murder` and Rape.
- Included is also the percent of the population in each state living in urban areas
- PC score vectors have length $n = 50$
- PC loading vectors have length $p = 4$
- PCA performed after standardizing each variable

Example: USA Arrests (ISLR)

- For each state in the US:
  - number of arrests per 100 000 residents for `Assault`, `Murder` and Rape.
- Included is also the percent of the population in each state living in urban areas
- PC score vectors have length $n = 50$
- PC loading vectors have length $p = 4$
- PCA performed after standardizing each variable

Example: USA Arrests Biplot

	PC1	PC2
Murder	0.5358995	−0.4181809
Assault	0.5831836	−0.1879856
UrbanPop	0.2781909	0.8728062
Rape	0.5434321	0.1673186



- **PC1**

  High loadings for Murder (0.536), Assault (0.583), and Rape (0.543):
  - These three variables contribute strongly and approximately equally to PC1.
  - PC1 could represent a general “crime severity” axis, as it captures patterns where these types of crimes tend to vary together.

  UrbanPop (0.279) has a smaller contribution:
  - Population density has less influence on PC1 compared to the crime-related variables.

- **PC2**

  High loading for UrbanPop (0.873):
  - PC2 is primarily influenced by UrbanPop.
  - This suggests PC2 captures variation in population density that is independent of crime severity.

  Negative contributions from Murder (-0.418) and Assault (-0.188):
  - Murder and Assault negatively influence PC2, indicating areas with high UrbanPop might have slightly lower relative crime rates.