When plotted using principal components, the and samples separate along principal component 2.

What combination of variables explain the differences between the and the samples?

![Diagram showing principal component analysis (PCA) results.](image)

---

```
[c, s, ~, ~, exp] = pca(data)
```

PCA returns up to $k = \min(n,p)$ principal components.

Principal component 1 explains 35.36% of the total variance. Principal component 2 explains another 23.02% of the total variance.

The coefficients (or "loadings") connect the original variables and the principal components.

Variables are mean centered during PCA, so "low" samples are negative and "high" samples are positive.

Therefore, the samples are low in variables 1 & 3 and high in variable p. The samples are low in variable p and high in variables 1 & 3.