1. Section 1

1.1 a.

The vertebral column data was first read from the ARFF file, then split into classes for processing.

```
library(foreign)
vert <- read.arff("column_2C_weka.arff")
vert_split <- split(vert, vert[,"class"])

sapply(vert_split$Abnormal[0:6], mean)
sapply(vert_split$Abnormal[0:6], median)
sapply(vert_split$Abnormal[0:6], sd)
sapply(vert_split$Normal[0:6], mean)
sapply(vert_split$Normal[0:6], median)
sapply(vert_split$Normal[0:6], sd)
```

1.1.1 Abnormal Data

mean

pelvic_incidence	pelvic_tilt	lumbar_lordosis_angle
64.69256	19.79111	55.92537
sacral_slope	pelvic_radius	degree_spondylolisthesis
44.90145	115.07771	37.77771

standard deviation

pelvic_incidence	pelvic_tilt	lumbar_lordosis_angle
65.27489	18.79890	56.15000
sacral_slope	pelvic_radius	degree_spondylolisthesis
44.63960	115.65032	31.94652

median

pelvic_incidence	pelvic_tilt	lumbar_lordosis_angle
17.66213	10.51587	19.66947
sacral_slope	pelvic_radius	degree_spondylolisthesis
14.51556	14.09060	40.69674

1.1.2 Normal Data

mean

pelvic_incidence	pelvic_tilt	lumbar_lordosis_angle
0	0	0
sacral_slope	pelvic_radius	degree_spondylolisthesis
0	0	0

standard deviation

pelvic_incidence	pelvic_tilt	lumbar_lordosis_angle
0	0	0
sacral_slope	pelvic_radius	degree_spondylolisthesis
0	0	0

median

pelvic_incidence	pelvic_tilt	lumbar_lordosis_angle
0	0	0
sacral_slope	pelvic_radius	degree_spondylolisthesis
0	0	0

1.2 b.

```
library(foreign)
vert <- read.arff("column_2C_weka.arff")
```

Figure 1.1: Caption Here

1.3 c.

2. Section 2

2.1 a.

Generating 100 3-dimensional vectors from a normal disribution with a mean vector as [1 2 1] and a 3x3 covariance matrix as [4 0.8 -0.3; 0.8 2 0.6; -0.3 0.6 5]

```
mean <- c(1,2,1)
sigma <- matrix(c(4, 0.8, -0.3, 0.8, 2, 0.6, -0.3, 0.6, 5), 3,3)
```

$$mean <- \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$$

$$sigma <- \begin{bmatrix} 4 & 0.8 & -0.3 \\ 0.8 & 2 & 0.6 \\ -0.3 & 0.6 & 5 \end{bmatrix}$$

2.2 b.

Figure 2.1: Caption Here

2.3 c.

```
library(fields)
Euclidean <- rdist(x1, x2)
```

$$x1 <- \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

$$x2 <- \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

$$Euclidean <- \begin{bmatrix} 0 \end{bmatrix}$$

```
library(stats)
Mahalanobis<-mahalanobis(x,mean,cov)
```

$$x < \dashv [$$

$$mean < \dashv [$$

$$cov < \dashv [$$

$$Mahalanobis < \dashv [$$

3. Section 3

3.1 Eigenvalues & Eigenvectors

```
records <- read.table("five-dimensional-records.txt")
  mean <- colMeans(records)
cov <- cov(records)
```

$$mean < - [$$

$$cov < - [$$

$$Eigenvalues < - [$$

$$Eigenvectors < - [$$

Figure 3.1: Caption Here

3.2 reduced representation

$$PCA < \dagger [$$

3.3 Scatter Plot

Figure 3.2: Caption Here

3.4 Reconstruct Data

$$records <- \begin{bmatrix} 5700 & 12.8 & 2500 & 270 & 25000 \\ 1000 & 10.9 & 600 & 10 & 10000 \\ 3400 & 8.8 & 1000 & 10 & 9000 \\ 3800 & 13.6 & 1700 & 140 & 25000 \\ 4000 & 12.8 & 1600 & 140 & 25000 \\ 8200 & 8.3 & 2600 & 60 & 12000 \\ 1200 & 11.4 & 400 & 10 & 16000 \\ 9100 & 11.5 & 3300 & 60 & 14000 \\ 9900 & 12.5 & 3400 & 180 & 18000 \\ 9600 & 13.7 & 3600 & 390 & 25000 \\ 9600 & 9.6 & 3300 & 80 & 12000 \\ 9400 & 11.4 & 4000 & 100 & 13000 \end{bmatrix}$$

$$reconstructed <- [$$

$$sqareerr <- [$$

4. Section 4

$$Eigenvalues < \dagger [$$

$$Eigenvectors < \dagger [$$

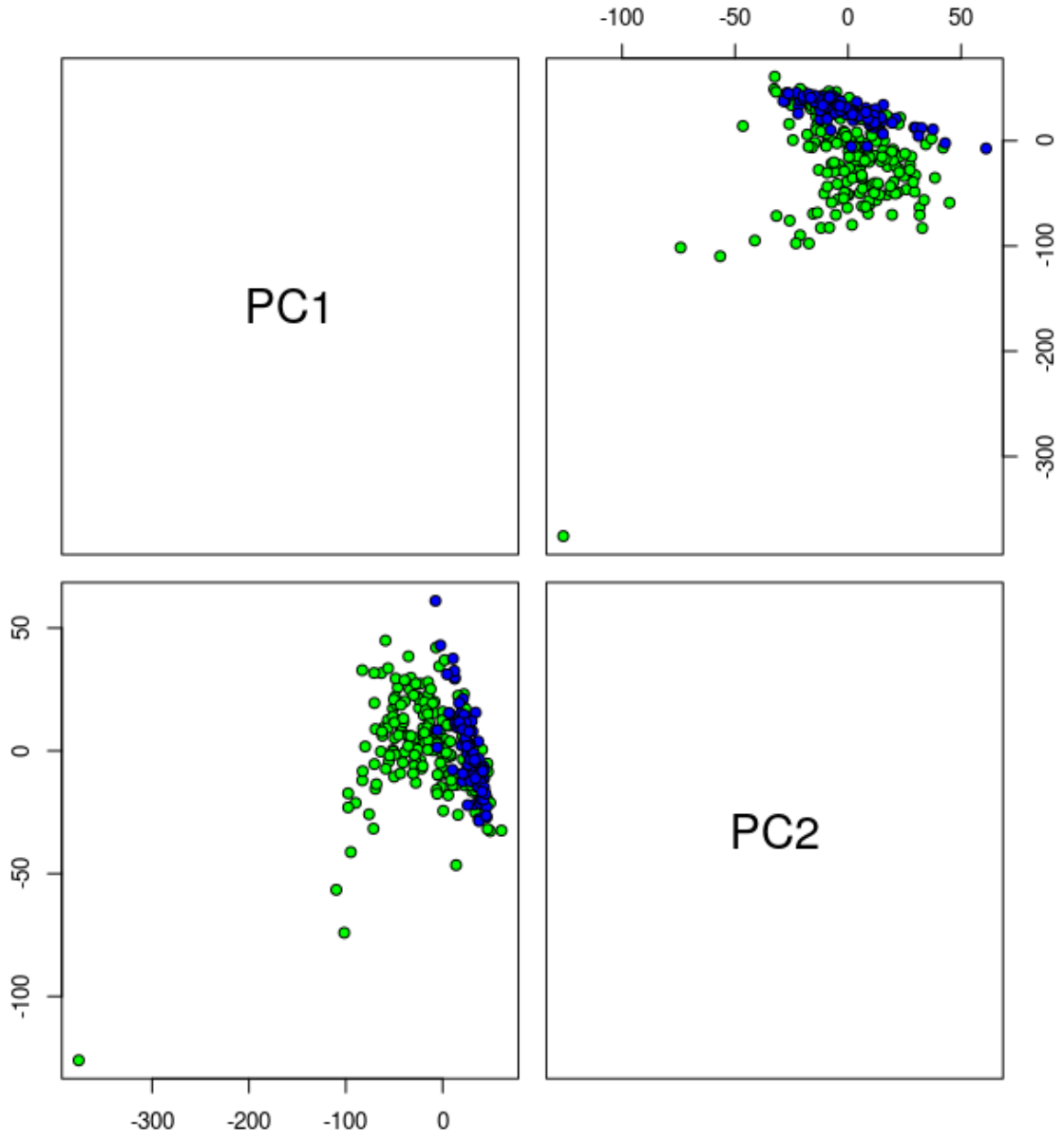

Figure 4.1: PCA Vertebral Column Data Set

5. Section 5

```
library('tsne')
records <- read.table('five-dimensional-records.txt')
tsne_10 <- tsne(records, perplexity=10)
tsne_50 <- tsne(records, perplexity=50)
```

$$tsne\_10 < \dashv [$$

$$tsne\_50 < \dashv [$$

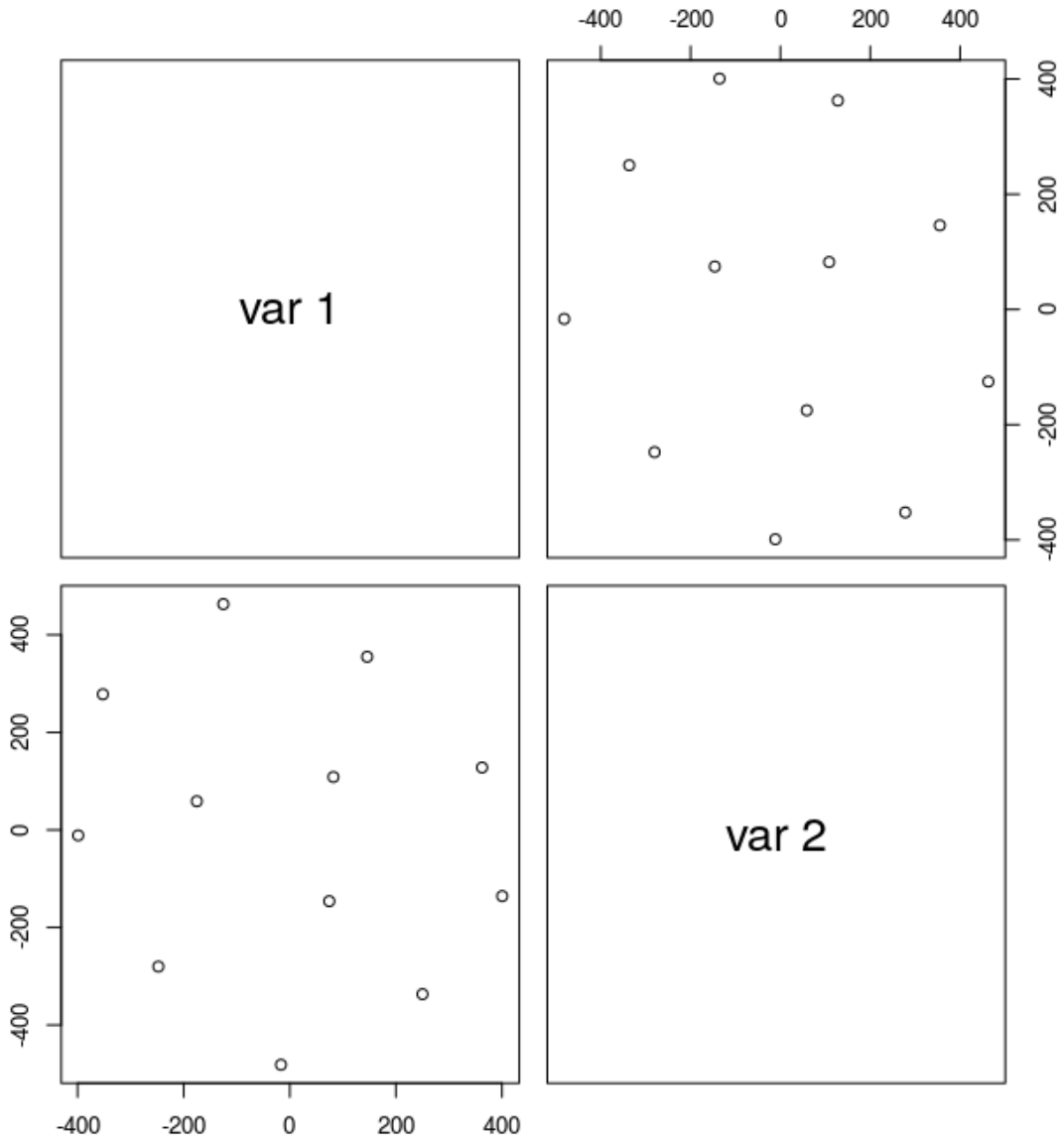

Figure 5.1: tsne perplexity 10

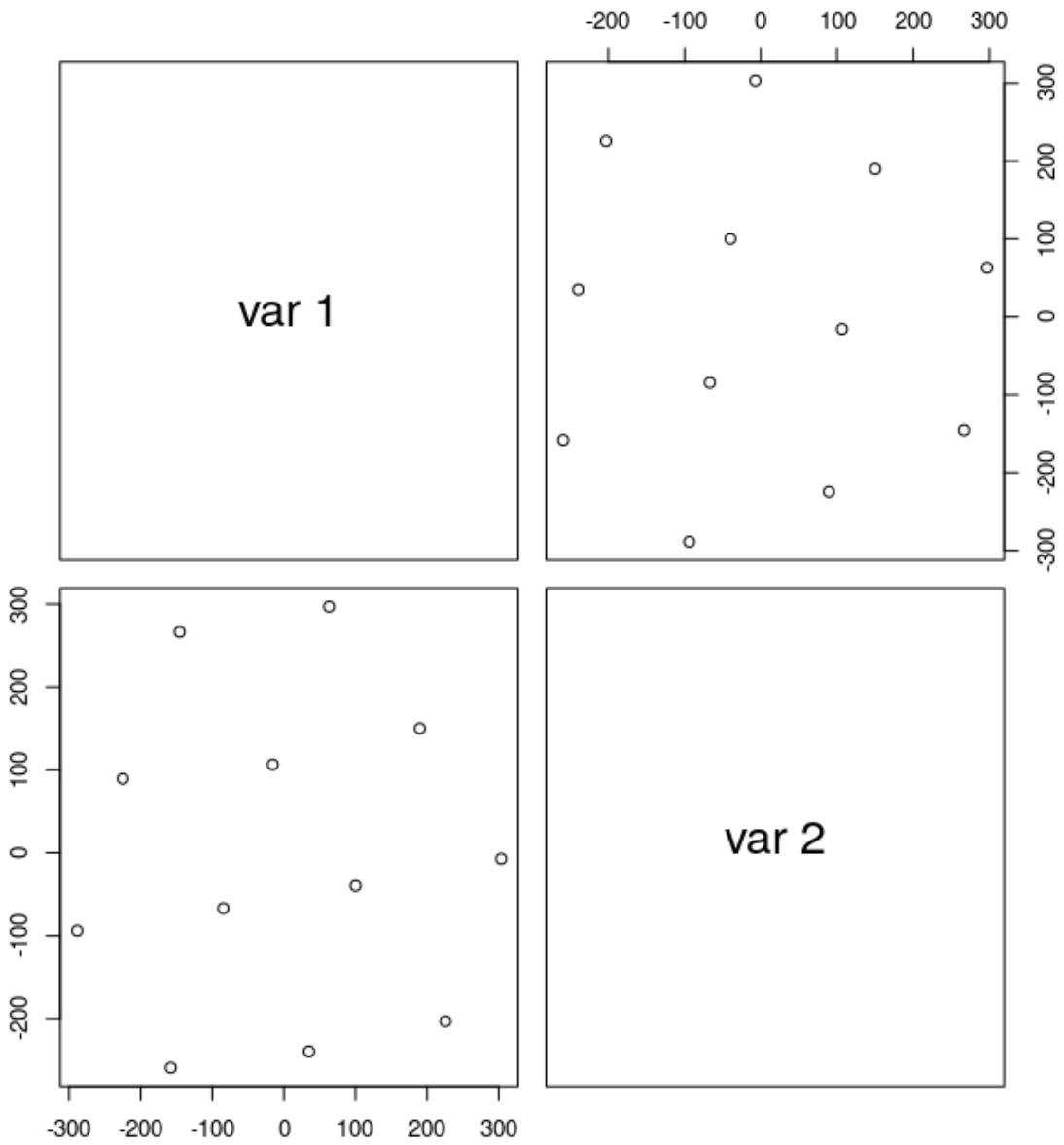

Figure 5.2: tsne perplexity 50