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
51.685244	12.821414	43.542605
sacral_slope	pelvic_radius	degree_spondylolisthesis
38.863830	123.890834	2.186572

standard deviation

pelvic_incidence	pelvic_tilt	lumbar_lordosis_angle
50.12312	13.48243	42.63892
sacral_slope	pelvic_radius	degree_spondylolisthesis
37.05969	123.87433	1.15271

median

pelvic_incidence	pelvic_tilt	lumbar_lordosis_angle
12.368161	6.778503	12.361388
sacral_slope	pelvic_radius	degree_spondylolisthesis
9.624004	9.014246	6.307483

1.2 b.

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

pairs(vert[0:6], pch = 21, bg = c(``green'', ``blue'')[unclass(vert$class)])
```

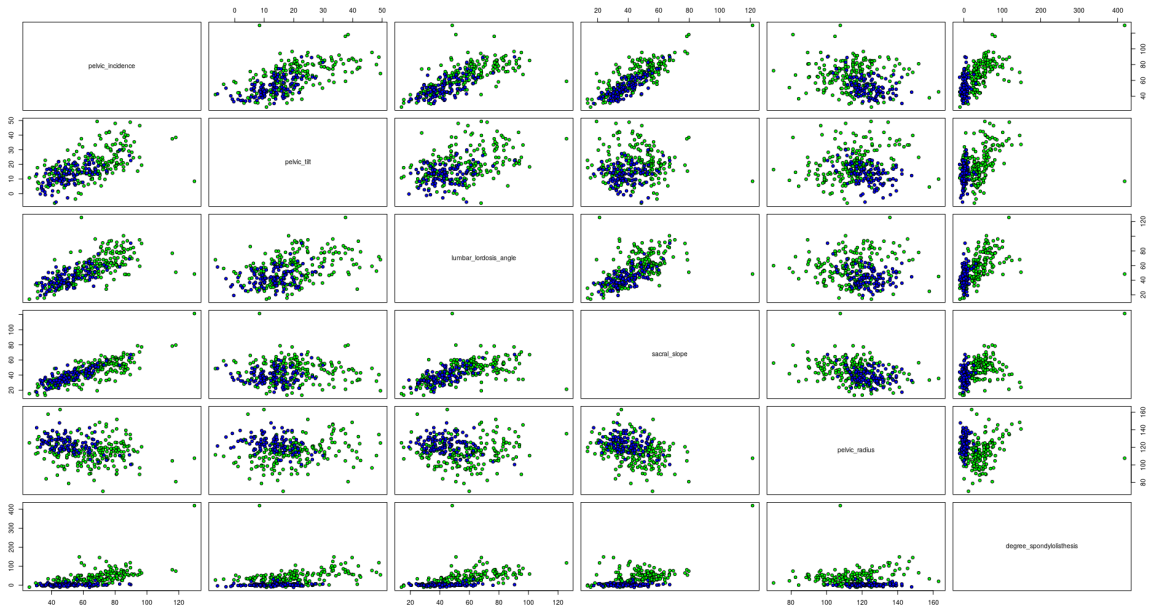

Figure 1.1: Feature Scatter Plot

1.3 c.

Given the values from section a and the scatter plot from section b we can see that the two classes are seperatated well by there values. For example if we pick two classes, pelvic_radius and degree_spondylolisthesis, we can compare the values and see how well they are seperated. If we also take into account the scatter plot from Figure 1.1 we can see that abnormal classes have a larger value with respect of degree_spondylolisthesis then the normal class.

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)
cov <- matrix(c(4, 0.8, -0.3, 0.8, 2, 0.6, -0.3, 0.6, 5), 3,3)
mvnd <- MASS\dotsmvrnorm(n = 100, mean, cov)
```

$$cov = \begin{bmatrix} 4 & 0.8 & -0.3 \\ 0.8 & 2 & 0.6 \\ -0.3 & 0.6 & 5 \end{bmatrix}$$

2.2 b.

2.3 c.

3. Section 3

4. Section 4

5. Section 5