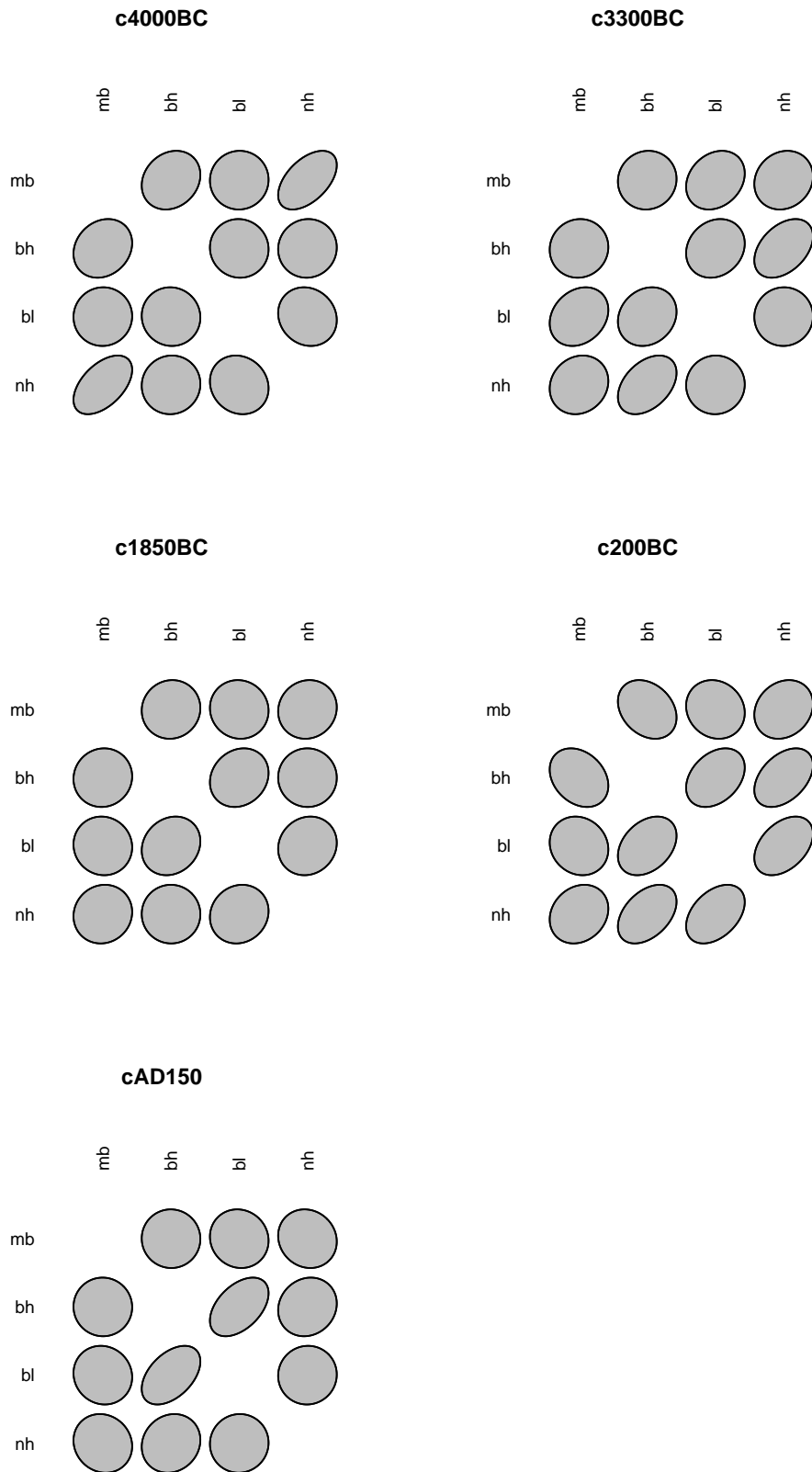


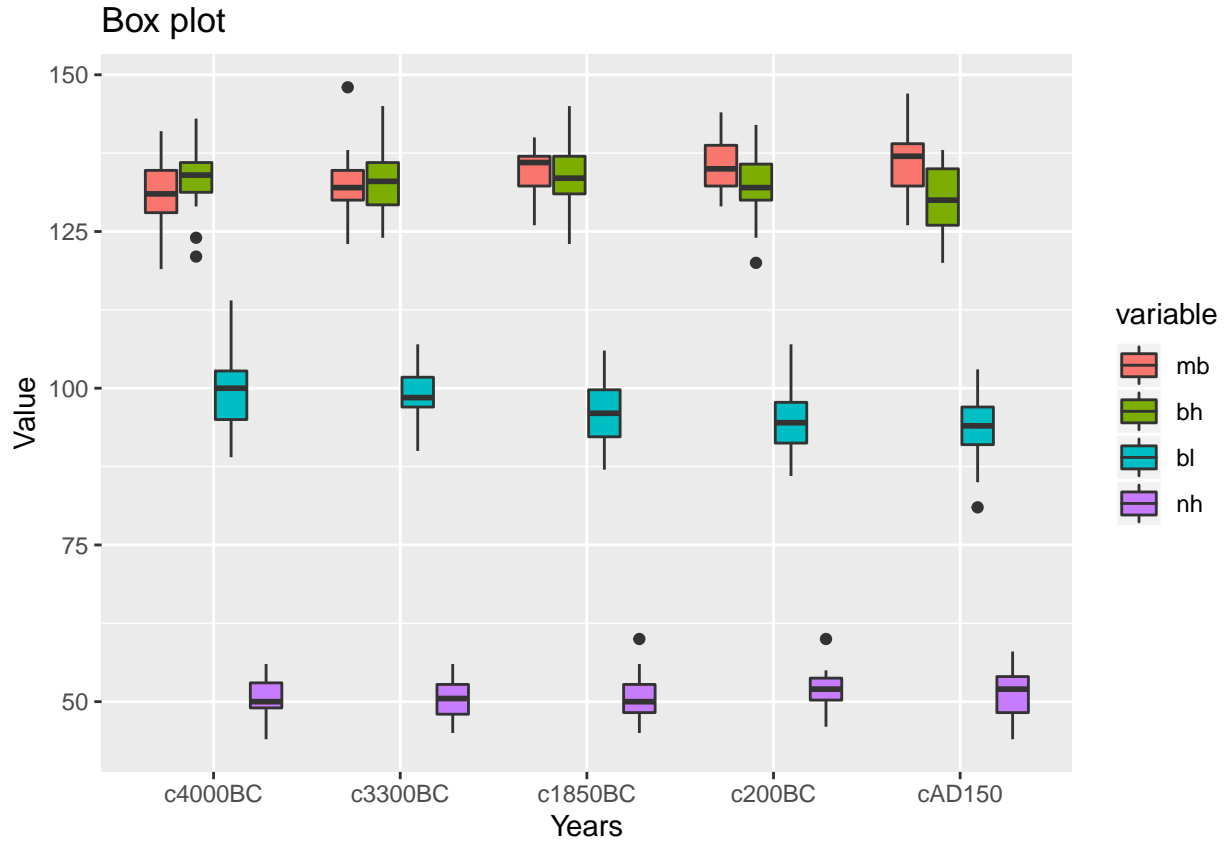
03

3 Comparison of mean vectors (one{way MANOVA)

a



From above correlation plot over 5 time period the common thing which we noticed is the positive correlation between basibregmatic height and basialiveolar length. Moreover, the correlation become more stronger between them over the period. In AD 150 the correlation between nh and mb was negative which was unprecedented.



From this the only thing which is informative is that the mean bl length decreased over the years while rest of the mean measurement did not follow any pattern. Also the distance between mean bh and mb increased towards the end of AD150. As maximal breadth of the skull increased, the height of the skull increased as well. Were brains were increasing in size?

b

Table 1: Mean of covariates in different groups

Group	mb	bh	bl	nh
c4000BC	131.3667	133.6000	99.16667	50.53333
c3300BC	132.3667	132.7000	99.06667	50.23333
c1850BC	134.4667	133.8000	96.03333	50.56667
c200BC	135.5000	132.3000	94.53333	51.96667
cAD150	136.1667	130.3333	93.50000	51.36667

```
manova_fit <- manova(cbind(mb, bh, bl, nh) ~ sk$epoch, sk)
summary(manova_fit, test="Pillai")
```

```
##          Df Pillai approx F num Df den Df    Pr(>F)
## sk$epoch   4 0.35331    3.512    16   580 4.675e-06 ***
```

```
## Residuals 145
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Pillai's trace is used as a test statistic in MANOVA. It is a positive valued statistic ranging from 0 to 1. Increasing values means that effects are contributing more to the model; you should reject the null hypothesis for large values.

Pillai's trace is considered to be the most powerful and robust statistic for general use, especially for departures from assumptions.

Here the p value is less than 0.05 so we reject the null hypothesis which is ,that the vector means are same across group.

c

```
t2_interval <- function(data, conf) {
n <- nrow(data)
p <- ncol(data)
x_ <- colMeans(data)
S <- cov(data)
offset <- sqrt(p * (n - 1) * qf(1 - conf, df1=p, df2=n - p) / (n - p) * diag(S) / n)
rbind(x_ - offset, x_ + offset)
}

conf_interval <- t2_interval(sk[,2:5],0.05)
```

```
kable(conf_interval,caption = "95% confidence interval") %>%
  kable_styling(latex_options = "hold")
```

```
\begin{table}[!h]
```

```
\caption{95% confidence interval}
```

mb	bh	bl	nh
132.7147	131.2755	95.076	50.10776
135.2320	133.8178	97.844	51.75890

```
\end{table}
```

```
## Warning: Ignoring unknown parameters: bin
```

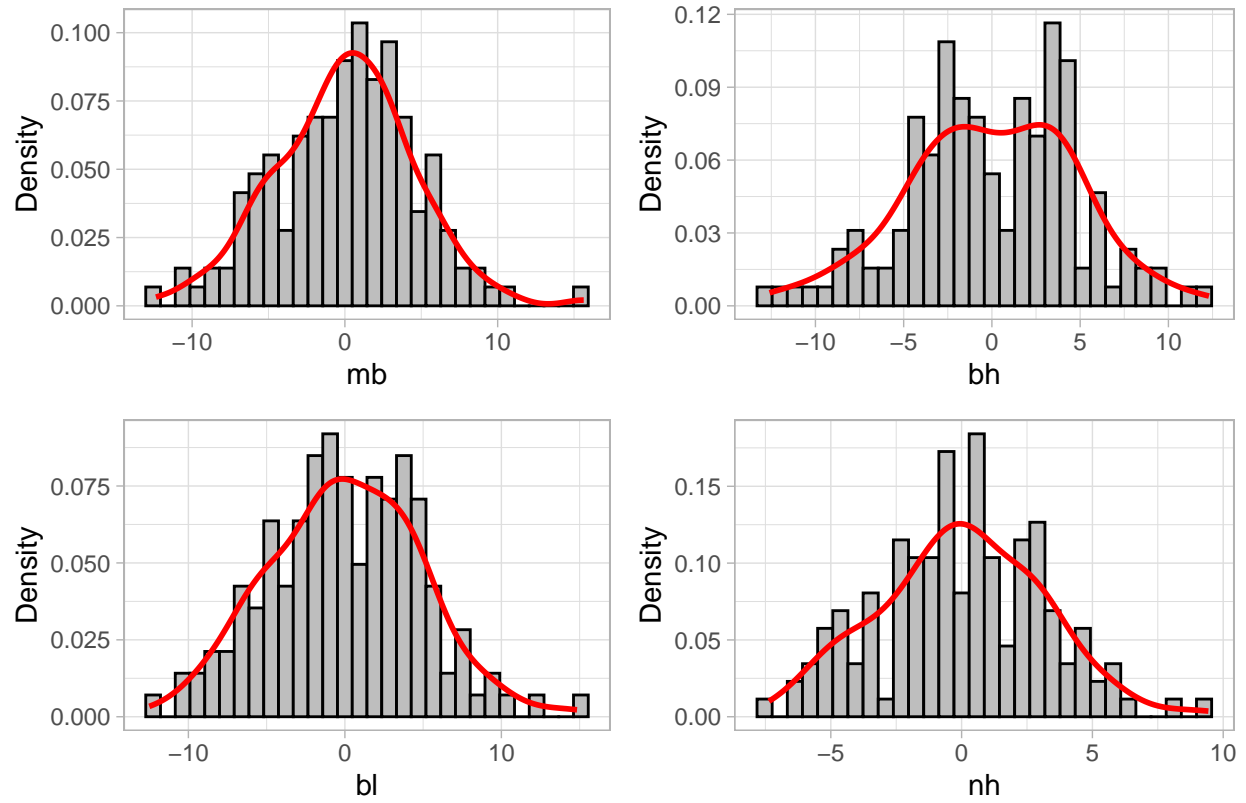
```
## Warning: Ignoring unknown parameters: bin
```

```
## Warning: Ignoring unknown parameters: bin
```

```
## Warning: Ignoring unknown parameters: bin
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

Density plot and histogram



The variables are nearly normally distributed.