

تحلیل مولفه اصلی اندازه استخوان آرواره

فراخوانی و تعریف داده‌ها :

```
> Ramus <- read.csv("C:/Users/12345/Desktop/Ramus.csv",header = T)
```

نمایش ۵ سطر ابتدایی داده‌ها :

```
> head(Ramus,5)
  year.8 year.8.5 year.9 year.9.5
1   47.8    48.8   49.0    49.7
2   46.4    47.3   47.7    48.4
3   46.3    46.8   47.8    48.5
4   45.1    45.3   46.1    47.2
5   47.6    48.5   48.9    49.3
```

محاسبه بردار میانگین :

```
> colMeans(Ramus)
  year.8 year.8.5   year.9 year.9.5
48.655  49.625   50.570  51.445
```

ماتریس واریانس-کواریانس (S) :

```
> (S<-cov(Ramus))
      year.8 year.8.5   year.9 year.9.5
year.8  6.329974 6.189079 5.777000 5.534763
year.8.5 6.189079 6.449342 6.153421 5.905658
year.9   5.777000 6.153421 6.918000 6.926684
year.9.5 5.534763 5.905658 6.926684 7.433132
```

مقادیر ویژه ماتریس S :

```
> (eigen<-eigen(S))
eigen() decomposition
$values
[1] 25.0546882  1.7433767  0.2185841  0.1137983

$vectors
      [,1]      [,2]      [,3]      [,4]
[1,] -0.4744413  0.5917697  0.5478640  0.3529293
[2,] -0.4922786  0.4055451 -0.4725736 -0.6081687
[3,] -0.5154434 -0.3035323 -0.5143660  0.6145030
[4,] -0.5166095 -0.6270646  0.4603769 -0.3577119

> landa<-eigen$values
```

کاهش بعد به کمک مقادیر ویژه :

```
> which(landa>mean(landa))  
[1] 1
```

در تحلیل مولفه اصلی با استفاده از ماتریس S مولفه‌ای که مقدار ویژه‌اش بزرگ تر از میانگین نقادیر ویژه باشد را نگه میداریم و بقیه را حذف میکنیم پس در اینجا به ۱ بعد کاهش میدهیم.

تحلیل مولفه اصلی

```
> Ramus.pca<-princomp(Ramus)  
> summary(Ramus.pca)  
Importance of components:  


|                        | Comp.1    | Comp.2     | Comp.3      | Comp.4      |
|------------------------|-----------|------------|-------------|-------------|
| Standard deviation     | 4.8787246 | 1.28693742 | 0.455691682 | 0.328798463 |
| Proportion of Variance | 0.9234897 | 0.06425905 | 0.008056783 | 0.004194488 |
| Cumulative Proportion  | 0.9234897 | 0.98774873 | 0.995805512 | 1.000000000 |

  
> loadings(Ramus.pca)
```

```
Loadings:  


|          | Comp.1 | Comp.2 | Comp.3 | Comp.4 |
|----------|--------|--------|--------|--------|
| year.8   | 0.474  | 0.592  | 0.548  | 0.353  |
| year.8.5 | 0.492  | 0.406  | -0.473 | -0.608 |
| year.9   | 0.515  | -0.304 | -0.514 | 0.615  |
| year.9.5 | 0.517  | -0.627 | 0.460  | -0.358 |


|                | Comp.1 | Comp.2 | Comp.3 | Comp.4 |
|----------------|--------|--------|--------|--------|
| SS loadings    | 1.00   | 1.00   | 1.00   | 1.00   |
| Proportion Var | 0.25   | 0.25   | 0.25   | 0.25   |
| Cumulative Var | 0.25   | 0.50   | 0.75   | 1.00   |


```

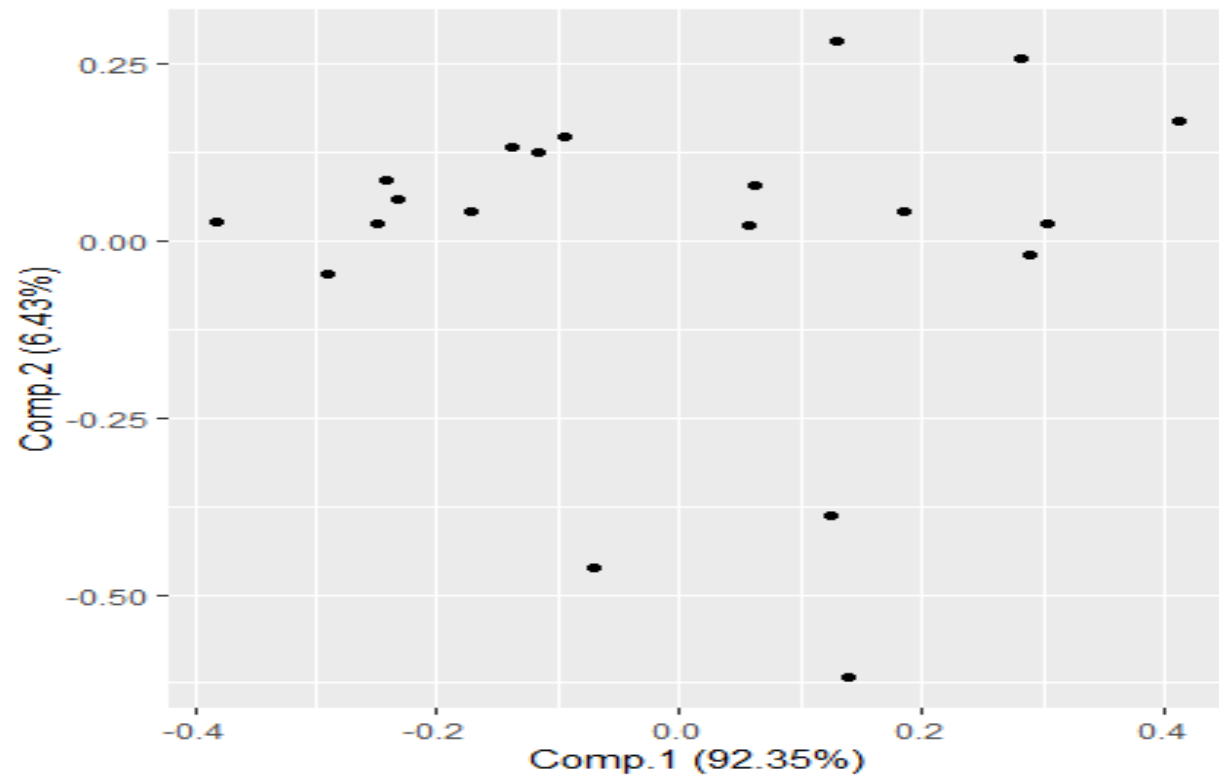
میبینیم که مولفه اصلی اول به تنهایی ۹۲ درصد از واریانس کل را نمایش میدهد پس در اینجا نیازی به بقیه مولفه ها نداریم و میتوانیم حذفشان کنیم.

رسم نمودار :

```
> require(ggfortify)  
> autoplot(Ramus.pca)
```

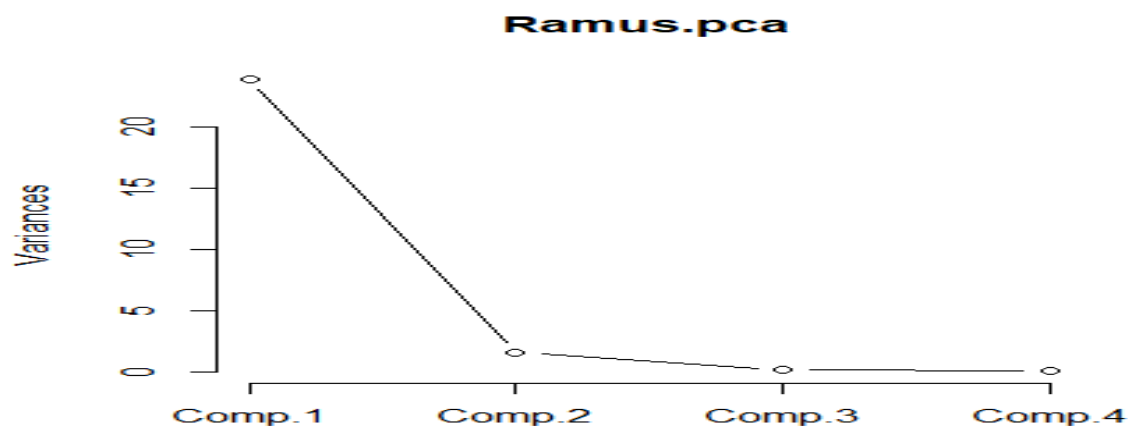
در نمودار زیر انحراف از توزیع نرمال مشاهده می‌شود چون داده‌ها حول نقطه صفر به خوبی پراکنده نشده اند.

همچنین به نظر میرسد ما دارای ۳ داده پرت هستیم



نمودار سنگ ریزه :

```
> plot(Ramus.pca,type = "lines")
```



در نمودار سنگ ریزه نیز نتایج بالا مبنی بر کافی بودن مولفه اصلی اول و حذف بقیه مولفه ها مشاهده می شود.

تحلیل مولفه اصلی با استفاده از ماتریس همبستگی (R)

```
> (R<-cor(Ramus))
      year.8 year.8.5 year.9 year.9.5
year.8  1.0000000 0.9686511 0.8729938 0.8068860
year.8.5 0.9686511 1.0000000 0.9212312 0.8529521
year.9   0.8729938 0.9212312 1.0000000 0.9659379
year.9.5 0.8068860 0.8529521 0.9659379 1.0000000
>
> (eigen<-eigen(R))
eigen() decomposition
$values
[1] 3.69524710 0.25534580 0.03256133 0.01684576

$vectors
      [,1]      [,2]      [,3]      [,4]
[1,] -0.4937453  0.5843508  0.5627874 -0.3130814
[2,] -0.5066122  0.3816454 -0.5250456  0.5674662
[3,] -0.5088771 -0.3386797 -0.4550180 -0.6475329
[4,] -0.4905138 -0.6310127  0.4478348  0.4008283

> landa<-eigen$values
>
> which(landa>1)
[1] 1
>
> Ramus.pca.R<-princomp(covmat=R)
> summary(Ramus.pca.R)
Importance of components:
              Comp.1      Comp.2      Comp.3      Comp.4
Standard deviation    1.9223026  0.50531752  0.180447589  0.129791235
Proportion of Variance 0.9238118  0.06383645  0.008140333  0.004211441
Cumulative Proportion 0.9238118  0.98764823  0.995788559  1.000000000
> loadings(Ramus.pca.R)

Loadings:
      Comp.1 Comp.2 Comp.3 Comp.4
year.8    0.494  0.584  0.563  0.313
year.8.5   0.507  0.382 -0.525 -0.567
year.9     0.509 -0.339 -0.455  0.648
year.9.5   0.491 -0.631  0.448 -0.401

      Comp.1 Comp.2 Comp.3 Comp.4
SS loadings    1.00    1.00    1.00    1.00
Proportion Var  0.25    0.25    0.25    0.25
Cumulative Var  0.25    0.50    0.75    1.00
```

با استفاده از ماتریس R باز به نتایج مشابه با قبل میرسیم با این تفاوت که میزان سهم مولفه اول در بیان واریانس کل به میزان کم تغییر میکند.