Class 8: Breast Cancer Mini Project

AUTHOR

Lisa chen (PID A17082974)

Before we get stuck into project work we will have a quick look at applying PCA to some example RNA Seq data (tail end of lab &)

Read the document

```
url2 <- "https://tinyurl.com/expression-CSV"
rna.data <- read.csv(url2, row.names=1)
head(rna.data)</pre>
```

```
      wt1
      wt2
      wt3
      wt4
      wt5
      ko1
      ko2
      ko3
      ko4
      ko5

      gene1
      439
      458
      408
      429
      420
      90
      88
      86
      90
      93

      gene2
      219
      200
      204
      210
      187
      427
      423
      434
      433
      426

      gene3
      1006
      989
      1030
      1017
      973
      252
      237
      238
      226
      210

      gene4
      783
      792
      829
      856
      760
      849
      856
      835
      885
      894

      gene5
      181
      249
      204
      244
      225
      277
      305
      272
      270
      279

      gene6
      460
      502
      491
      491
      493
      612
      594
      577
      618
      638
```

Q1. How many genes are in this data set? nrow (rna. data)

```
nrow(rna.data)
```

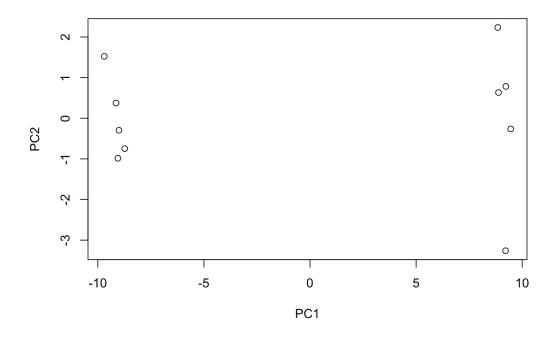
[1] 100

generating barplot etc. to make sense of this data is really not an exciting or worthwhile option to consider. So lets do PCA and plot the results:

##Run PCA

```
# Again we have to take the transpose of our data
pca <- prcomp(t(rna.data), scale =TRUE)

##Simple un polished plot of pc1 and pc2
plot(pca$x[,1], pca$x [,2], xlab="PC1", ylab="PC2")</pre>
```



take summary

```
summary(pca)
```

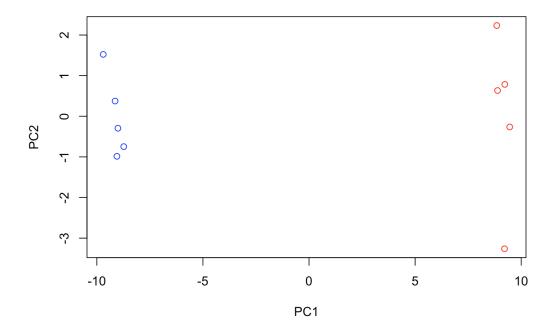
Importance of components:

PC1 PC2 PC3 PC4 PC5
PC6 PC7
Standard deviation 9.6237 1.5198 1.05787 1.05203 0.88062
0.82545 0.80111
Proportion of Variance 0.9262 0.0231 0.01119 0.01107 0.00775
0.00681 0.00642
Cumulative Proportion 0.9262 0.9493 0.96045 0.97152 0.97928
0.98609 0.99251

A quick barplot summary of this Proportion of variance for each PC can be obtained by calling the plot() function directly on our prcomp result object.

```
pca$x
```

```
PC1
                    PC2
                              PC3
                                                   PC5
                                        PC4
PC6
wt1 -9.697374 1.5233313 -0.2753567 0.7322391 -0.6749398
1.1823860
wt2 -9.138950 0.3748504 1.0867958 -1.9461655 0.7571209
-0.4369228
wt3 -9.054263 -0.9855163 0.4152966 1.4166028 0.5835918
0.6937236
wt4 -8.731483 -0.7468371 0.5875748 0.2268129 -1.5404775
-1.2723618
wt5 -9.006312 -0.2945307 -1.8498101 -0.4303812 0.8666124
-0.2496025
ko1 8.846999 2.2345475 -0.1462750 -1.1544333 -0.6947862
0.7128021
ko2 9.213885 -3.2607503 0.2287292 -0.7658122 -0.4922849
0.9170241
ko3 9.458412 -0.2636283 -1.5778183 0.2433549 0.3654124
-0.5837724
ko4 8.883412 0.6339701 1.5205064 0.7760158 1.2158376
-0.1446094
ko5 9.225673 0.7845635 0.0103574 0.9017667 -0.3860869
-0.8186668
           PC7
                       PC8
                                  PC9
                                             PC10
wt1 -0.24446614 1.03519396 0.07010231 3.388516e-15
wt2 -0.03275370 0.26622249 0.72780448 2.996563e-15
wt3 -0.03578383 -1.05851494 0.52979799 3.329630e-15
wt4 -0.52795595 -0.20995085 -0.50325679 3.317526e-15
wt5 0.83227047 -0.05891489 -0.81258430 2.712504e-15
ko1 -0.07864392 -0.94652648 -0.24613776 2.768138e-15
ko2 0.30945771 0.33231138 -0.08786782 3.317091e-15
ko3 -1.43723425 0.14495188 0.56617746 3.299214e-15
ko5 1.56584821 0.19140827 0.62950330 2.785473e-15
# we have 5 wt and 5 ko samples
mycols <- c(rep("blue", 5), rep("red", 5))</pre>
mycols
 [1] "blue" "blue" "blue" "blue" "red" "red" "red"
"red" "red"
 plot(pca$x[,1], pca$x [,2], xlab="PC1", ylab="PC2", col=mycols)
```



I could examine which genes contribute most to this first pt

```
head(sort(abs(pca$rotation[,1]), decreasing = T))
```

gene100 gene66 gene45 gene68 gene98 gene60 0.1038708 0.1038455 0.1038402 0.1038395 0.1038372 0.1038055

#Analysis of Breast Cancer FNA data.

The data itself comes from Wisconsin Breast Cancer Diagnostic Data Set.

Values in this data set describe cahracteristics of the cell nuclei present in digitized images. of a fine needle aspiration (FNA) of a breast mass.

```
wisc.df <- read.csv("WisconsinCancer (3).csv", row.names=1)
head(wisc.df)</pre>
```

	uragnosis	rautus_iiiean	texture_mean	per fille cer _illean
area_mear	1			
842302	M	17.99	10.38	122.80
1001.0				
842517	М	20.57	17.77	132.90
1326.0				
84300903	М	19.69	21.25	130.00

1203.0	N 4	11 12	20.20	77.50
84348301	М	11.42	20.38	77.58
386.1	N4	20.20	14 24	125 10
84358402	М	20.29	14.34	135.10
1297.0		12 45	15 70	02.57
843786	М	12.45	15.70	82.57
477.1				
	_	ean compa	ctness_mean con	cavity_mean
•	oints_mean 0.11	0.40	0.27760	0 2001
842302	0.11	540	0.27700	0.3001
0.14710	0.00	474	0 07064	0.0060
842517	0.08	4/4	0.07864	0.0869
0.07017	0 10	0.00	0 15000	0 1074
84300903	0.10	900	0.15990	0.1974
0.12790	0 14	250	0 20200	0 2414
84348301	0.14	250	0.28390	0.2414
0.10520	0.10	220	0 12200	0.1000
84358402	0.10	030	0.13280	0.1980
0.10430	0.10	700	0 17000	0.4570
843786	0.12	780	0.17000	0.1578
0.08089		. £1	4:	
	symmetry_mea e perimeter_:	_	_dimension_mean	radius_se
842302	e perimeter_: 0.241		0.07871	1.0950
0.9053	8.589	9	0.0/8/1	1.0950
842517	0.181	2	0.05667	0.5435
0.7339	3.398	<u>Z</u>	0.03007	0.3433
84300903	0.206	n	0.05999	0.7456
0.7869	4.585	9	0.03999	0.7430
84348301	0. 259	7	0.09744	0.4956
1.1560		<i>I</i>	0.09/44	0.4930
84358402		o	0.05883	0.7572
0.7813	5.438	9	0.03003	01/3/2
843786	0.208	7	0.07613	0.3345
0.8902	2.217	,	0.07013	0.3343
		thnacc ca	compactness_se	concavity se
concave.p	-	tille33_3e	compactifess_se	concavity_se
842302	153.40	0.006399	0.04904	0.05373
0.01587	133.40	0.000333	0.04304	0.05575
842517	74.08	0.005225	0.01308	0.01860
0.01340	74.00	0.003223	0.01300	0.01000
84300903	94_03	0.006150	0.04006	0.03832
0.02058	57105	0.000130	0104000	0103032
84348301	27.23	0.009110	0.07458	0.05661
0.01867	21 1 23	0.003110	010/430	0.0001
0.0100/				

84358402 0.01885	94.44 0.	011490	0.02461	0.05688
843786 0.01137	27.19 0.	007510	0.03345	0.03672
	symmetry_se fra	ctal dimens:	ion co rad	ius vorst
texture_w	, , , –	ctat_uimens.	LOII_SE Tau.	Tus_worst
842302	0.03003	0 (006193	25.38
042302 17.33	0.03003	0.0	000193	23.30
842517	0.01389	ο (003532	24.99
23.41	0.01309	0.0	003332	24.99
84300903	0.02250	0 (004571	23.57
25.53	0.02230	0.0	004371	23.37
84348301	0.05963	0 (009208	14.91
26.50	0.03903	0.0	009200	14.91
84358402	0.01756	0 (005115	22.54
16.67	0.01750	0.0	003113	22 . 34
843786	0.02165	0 0	005082	15.47
23.75	0.02103	0.0	003002	13.47
	oerimeter_worst	area worst	smoothness	s worst
compactnes	-	arca_worse	Silloo crimes.	3_w0130
842302	184.60	2019.0		0.1622
0.6656	101100	2013.0		0.1022
842517	158.80	1956.0		0.1238
0.1866	150100	100010		0.1200
84300903	152.50	1709.0		0.1444
0.4245				
84348301	98.87	567.7		0.2098
0.8663				
84358402	152.20	1575.0		0.1374
0.2050				
843786	103.40	741.6		0.1791
0.5249				
(concavity_worst	concave.po	ints_worst	symmetry_worst
842302	0.7119		0.2654	0.4601
842517	0.2416		0.1860	0.2750
84300903	0.4504		0.2430	0.3613
84348301	0.6869		0.2575	0.6638
84358402	0.4000		0.1625	0.2364
843786	0.5355		0.1741	0.3985
-	fractal_dimensi	on_worst		
842302		0.11890		
842517		0.08902		
84300903		0.08758		
84348301		0.17300		
84358402		0.07678		

843786 0.12440

```
diagnosis <- as.factor(wisc.df$diagnosis)</pre>
```

Now I want to make sure I remove that column from my dataset for analysis

```
wisc.data <- wisc.df [,-1]
head(wisc.data)</pre>
```

	radius_mean	texture_mea	n perimete	er_mean	area_mean
smoothnes	-	_	·	_	_
842302	17.99	10.3	8	122.80	1001.0
0.11840					
842517	20.57	17.7	7	132.90	1326.0
0.08474					
84300903	19.69	21.2	5	130.00	1203.0
0.10960					
84348301	11.42	20.3	8	77.58	386.1
0.14250					
84358402	20.29	14.3	4	135.10	1297.0
0.10030					
843786	12.45	15.7	0	82.57	477.1
0.12780					
	compactness	_mean concav	ity_mean d	concave.	points_mean
symmetry_	_mean				
842302	0.2	27760	0.3001		0.14710
0.2419					
842517	0.0	07864	0.0869		0.07017
0.1812					
84300903	0.1	15990	0.1974		0.12790
0.2069					
84348301	0.2	28390	0.2414		0.10520
0.2597					
84358402	0.1	13280	0.1980		0.10430
0.1809					
843786	0.1	17000	0.1578		0.08089
0.2087					
		ension_mean	radius_se	texture	e_se
•	_se area_se				
842302		0.07871	1.0950	0.9	9053
	53.40				
842517		0.05667	0.5435	0.7	7339

3.398	74.08	3			
8430090	3		0.05999	0.7456	0.7869
4.585	94.03	3			
8434830	1		0.09744	0.4956	1.1560
3.445	27.23	3			
8435840	2		0.05883	0.7572	0.7813
5.438	94.44	ļ			
843786			0.07613	0.3345	0.8902
2.217	27.19)			
	smoo	thness_se o	compactness_	_se conc	avity_se
concave	<pre>.point</pre>	s_se			
842302		0.006399	0.049	904	0.05373
0.01587					
842517		0.005225	0.013	308	0.01860
0.01340					
8430090	3	0.006150	0.04	006	0.03832
0.02058					
8434830	1	0.009110	0.07	458	0.05661
0.01867					
8435840	2	0.011490	0.02	461	0.05688
0.01885					
843786		0.007510	0.033	345	0.03672
0.01137					
		netry_se fra	actal_dimen	sion_se	radius_worst
texture	symm		actal_dimen	sion_se	radius_worst
	symn _worst		_	sion_se	radius_worst 25.38
texture 842302 17.33	symm _worst	0.03003	_	_	_
texture 842302 17.33 842517	symm _worst	<u></u>	0	_	_
texture 842302 17.33 842517 23.41	symn _worst	0.03003 0.01389	0	.006193 .003532	25.38 24.99
texture 842302 17.33 842517	symn _worst	0.03003	0	.006193 .003532	25.38
texture 842302 17.33 842517 23.41 8430090 25.53	symn _worst	0.03003 0.01389 0.02250	0	.006193 .003532 .004571	25.38 24.99 23.57
texture 842302 17.33 842517 23.41 8430090	symn _worst	0.03003 0.01389	0	.006193 .003532	25.38 24.99
texture 842302 17.33 842517 23.41 8430090 25.53 8434830 26.50	symm _worst 3	0.03003 0.01389 0.02250 0.05963	0 0	.006193 .003532 .004571 .009208	25.38 24.99 23.57 14.91
texture 842302 17.33 842517 23.41 8430090 25.53 8434830 26.50 8435840	symm _worst 3	0.03003 0.01389 0.02250	0 0	.006193 .003532 .004571	25.38 24.99 23.57
texture 842302 17.33 842517 23.41 8430090 25.53 8434830 26.50 8435840 16.67	symm _worst 3	0.03003 0.01389 0.02250 0.05963	0 0	.006193 .003532 .004571 .009208	25.38 24.99 23.57 14.91
texture 842302 17.33 842517 23.41 8430090 25.53 8434830 26.50 8435840 16.67 843786	symm _worst 3 1	0.03003 0.01389 0.02250 0.05963	0 0 0	.006193 .003532 .004571 .009208	25.38 24.99 23.57 14.91 22.54
texture 842302 17.33 842517 23.41 8430090 25.53 8434830 26.50 8435840 16.67	symm _worst 3 1	0.03003 0.01389 0.02250 0.05963 0.01756 0.02165	0 0 0	.006193 .003532 .004571 .009208 .005115	25.38 24.99 23.57 14.91 22.54 15.47
texture 842302 17.33 842517 23.41 8430090 25.53 8434830 26.50 8435840 16.67 843786 23.75	symm _worst 3 1 2	0.03003 0.03003 0.01389 0.02250 0.05963 0.01756 0.02165	0 0 0	.006193 .003532 .004571 .009208 .005115	25.38 24.99 23.57 14.91 22.54 15.47
texture 842302 17.33 842517 23.41 8430090 25.53 8434830 26.50 8435840 16.67 843786 23.75 compact	symm _worst 3 1 2	0.03003 0.03003 0.01389 0.02250 0.05963 0.01756 0.02165 dimeter_worst	0 0 0 0 area_wors	.006193 .003532 .004571 .009208 .005115 .005082 t smooth	25.38 24.99 23.57 14.91 22.54 15.47 ness_worst
texture 842302 17.33 842517 23.41 8430090 25.53 8434830 26.50 8435840 16.67 843786 23.75 compact 842302	symm _worst 3 1 2	0.03003 0.03003 0.01389 0.02250 0.05963 0.01756 0.02165	0 0 0 0 area_wors	.006193 .003532 .004571 .009208 .005115 .005082 t smooth	25.38 24.99 23.57 14.91 22.54 15.47
texture 842302 17.33 842517 23.41 8430090 25.53 8434830 26.50 8435840 16.67 843786 23.75 compact 842302 0.6656	symm _worst 3 1 2	0.03003 0.03003 0.01389 0.02250 0.05963 0.01756 0.02165 Emeter_worst	0 0 0 0 area_wors	.006193 .003532 .004571 .009208 .005115 .005082 t smooth	25.38 24.99 23.57 14.91 22.54 15.47 ness_worst 0.1622
texture 842302 17.33 842517 23.41 8430090 25.53 8434830 26.50 8435840 16.67 843786 23.75 compact 842302 0.6656 842517	symm _worst 3 1 2	0.03003 0.03003 0.01389 0.02250 0.05963 0.01756 0.02165 dimeter_worst	0 0 0 0 area_wors	.006193 .003532 .004571 .009208 .005115 .005082 t smooth	25.38 24.99 23.57 14.91 22.54 15.47 ness_worst
texture 842302 17.33 842517 23.41 8430090 25.53 8434830 26.50 8435840 16.67 843786 23.75 compact 842302 0.6656 842517 0.1866	symm _worst 3 1 2 peri	0.03003 0.03003 0.01389 0.02250 0.05963 0.01756 0.02165 Emeter_worst vorst 184.60	0 0 0 0 area_wors 0 2019.0	.006193 .003532 .004571 .009208 .005115 .005082 t smooth	25.38 24.99 23.57 14.91 22.54 15.47 ness_worst 0.1622 0.1238
texture 842302 17.33 842517 23.41 8430090 25.53 8434830 26.50 8435840 16.67 843786 23.75 compact 842302 0.6656 842517	symm _worst 3 1 2 peri	0.03003 0.03003 0.01389 0.02250 0.05963 0.01756 0.02165 Emeter_worst	0 0 0 0 area_wors 0 2019.0	.006193 .003532 .004571 .009208 .005115 .005082 t smooth	25.38 24.99 23.57 14.91 22.54 15.47 ness_worst 0.1622

84348301	98.87	567.7		0.2098
0.8663				
84358402	152.20	1575.0		0.1374
0.2050				
843786	103.40	741.6		0.1791
0.5249				
	concavity_worst	concave.po	ints_worst	symmetry_worst
842302	0.7119		0.2654	0.4601
842517	0.2416		0.1860	0.2750
84300903	0.4504		0.2430	0.3613
84348301	0.6869		0.2575	0.6638
84358402	0.4000		0.1625	0.2364
843786	0.5355		0.1741	0.3985
	fractal_dimension	on_worst		
842302		0.11890		
842517		0.08902		
84300903		0.08758		
84348301		0.17300		
84358402		0.07678		
843786		0.12440		

Q1 How many observations are in this dataset

Q2 How many of the observations have a malignant diagnosis

table(wisc.df\$diagnosis)

B M 357 212

Q3. How many variables/features in the data are suffixed with _mean?

colnames (wisc.data)

```
[1] "radius_mean" "texture_mean"
[3] "perimeter_mean" "area_mean"
[5] "smoothness_mean" "compactness_mean"
[7] "concavity_mean" "concave.points_mean"
[9] "symmetry_mean" "fractal_dimension_mean"
```

```
[11] "radius_se"
                                "texture_se"
[13] "perimeter_se"
                                "area_se"
[15] "smoothness_se"
                                "compactness_se"
[17] "concavity_se"
                                "concave.points_se"
[19] "symmetry_se"
                                "fractal_dimension_se"
[21] "radius_worst"
                                "texture_worst"
[23] "perimeter_worst"
                                "area_worst"
[25] "smoothness_worst"
                                "compactness_worst"
[27] "concavity_worst"
                                "concave.points_worst"
[29] "symmetry_worst"
                                "fractal_dimension_worst"
```

```
length(grep("_mean", colnames(wisc.data), value=TRUE) )
```

[1] 10

#Principal component analysis

here we will use 'prcomp()' on the 'wisc.data' object - the one without the diagnosis column.

First, we have decide whether to use the 'scale=TRUE' argument when we run 'prcomp()'

We can look at the means and sd of each column. If they are similar then we are all good to go. If not we should use 'scale= TRUE'

colMeans(wisc.data)

radius_mean	texture_mean
perimeter_mean	
1.412729e+01	1.928965e+01
9.196903e+01	
area_mean	smoothness_mean
compactness_mean	
6.548891e+02	9.636028e-02
1.043410e-01	
concavity_mean	concave.points_mean
symmetry_mean	
8.879932e-02	4.891915e-02
1.811619e-01	
fractal_dimension_mean	radius_se
texture_se	
6.279761e-02	4.051721e-01

1.216853e+00	
perimeter_se	area_se
smoothness_se	
2.866059e+00	4.033708e+01
7.040979e-03	
compactness_se	concavity_se
concave.points_se	
2.547814e-02	3.189372e-02
1.179614e-02	
symmetry_se	<pre>fractal_dimension_se</pre>
radius_worst	
2.054230e-02	3.794904e-03
1.626919e+01	
texture_worst	perimeter_worst
area_worst	
2.567722e+01	1.072612e+02
8.805831e+02	
smoothness_worst	compactness_worst
concavity_worst	
1.323686e-01	2.542650e-01
2.721885e-01	
concave.points_worst	symmetry_worst
<pre>fractal_dimension_worst</pre>	
1.146062e-01	2.900756e-01
8.394582e-02	

apply(wisc.data, 2, sd)

radius_mean	texture_mean
perimeter_mean	
3.524049e+00	4.301036e+00
2.429898e+01	
area_mean	smoothness_mean
compactness_mean	
3.519141e+02	1.406413e-02
5.281276e-02	
concavity_mean	concave.points_mean
symmetry_mean	
7 . 971981e-02	3.880284e-02
2.741428e-02	
<pre>fractal_dimension_mean</pre>	radius_se
texture_se	
7 . 060363e-03	2.773127e-01
5.516484e-01	

http://localhost:7017/ Page 11 of 33

perimeter_se	area_se
smoothness_se	
2.021855e+00	4.549101e+01
3.002518e-03	
compactness_se	concavity_se
concave.points_se	
1.790818e-02	3.018606e-02
6.170285e-03	
symmetry_se	<pre>fractal_dimension_se</pre>
radius_worst	
8.266372e-03	2.646071e-03
4.833242e+00	
texture_worst	perimeter_worst
area_worst	
6.146258e+00	3.360254e+01
5.693570e+02	
smoothness_worst	compactness_worst
concavity_worst	
2.283243e-02	1.573365e-01
2.086243e-01	
concave.points_worst	symmetry_worst
<pre>fractal_dimension_worst</pre>	
6.573234e-02	6.186747e-02
1.806127e-02	

```
wisc.pr <- prcomp(wisc.data, scale=TRUE)
summary(wisc.pr)</pre>
```

Importance of components:

·	PC1	PC2	PC3	PC4	PC5
PC6 PC7	-	-			
Standard deviation	3.6444	2.3857	1.67867	1.40735	1.28403
1.09880 0.82172					
Proportion of Variance	0.4427	0.1897	0.09393	0.06602	0.05496
0.04025 0.02251					
Cumulative Proportion	0.4427	0.6324	0.72636	0.79239	0.84734
0.88759 0.91010					
	PC8	B PCS	PC10	PC11	PC12
PC13 PC14					
Standard deviation	0.69037	0.6457	0.59219	0.5421	0.51104
0.49128 0.39624					
Proportion of Variance	0.01589	0.0139	0.01169	0.0098	0.00871
0.00805 0.00523					
Cumulative Proportion	0.92598	0.9399	0.95157	0.9614	0.97007

http://localhost:7017/ Page 12 of 33

0.97812 0.98335

PC15 PC16 PC17 PC18 **PC19** PC20 PC21 Standard deviation 0.30681 0.28260 0.24372 0.22939 0.22244 0.17652 0.1731 Proportion of Variance 0.00314 0.00266 0.00198 0.00175 0.00165 0.00104 0.0010 Cumulative Proportion 0.98649 0.98915 0.99113 0.99288 0.99453 0.99557 0.9966 PC22 PC23 PC24 PC25 PC26 PC27 PC28 Standard deviation 0.16565 0.15602 0.1344 0.12442 0.09043 0.08307 0.03987 Proportion of Variance 0.00091 0.00081 0.0006 0.00052 0.00027 0.00023 0.00005 Cumulative Proportion 0.99749 0.99830 0.9989 0.99942 0.99969 0.99992 0.99997 PC29 PC30 Standard deviation 0.02736 0.01153 Proportion of Variance 0.00002 0.00000 Cumulative Proportion 1.00000 1.00000

Q4. From your results, what proportion of the original variance is captured by the first principal components (PC1)?

44.27%

Q5. How many principal components (PCs) are required to describe at least 70% of the original variance in the data?

3 PCs capture 72.6% of the original variance

Plotting the PCA results

#biplot(wisc.pr)

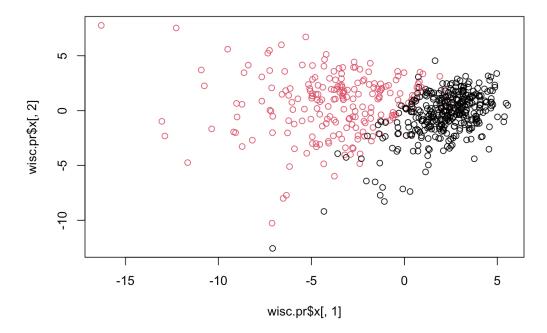
we need to make our own plot

attributes (wisc.pr)

http://localhost:7017/ Page 13 of 33

```
$names
[1] "sdev" "rotation" "center" "scale" "x"
$class
[1] "prcomp"
```

```
plot(wisc.pr$x[,1], wisc.pr$x[,2], col=diagnosis)
```

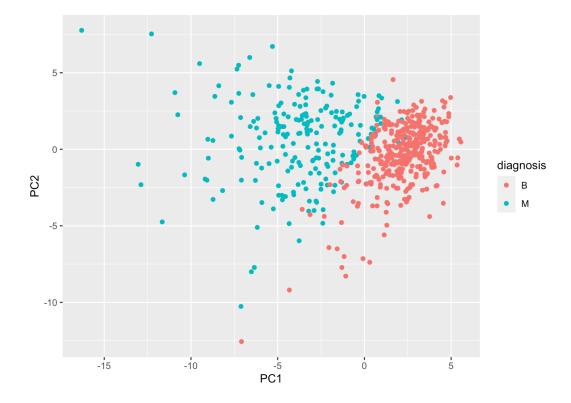


```
library(ggplot2)

pc <- as.data.frame(wisc.pr$x)

ggplot(pc) +
  aes(PC1, PC2, col=diagnosis) +
  geom_point()</pre>
```

Class 8: Breast Cancer Mini Project



Q6. How many principal components (PCs) are required to describe at least 90% of the original variance in the data?

7 PCS are required to describe 90% of the original variance.

Q7. What stands out to you about this plot? Is it easy or difficult to understand and why?

I like how the diagnoses are not interlapped, which shows distinguishable differences. It is however hard to understand.

Q9. For the first principal component, what is the component of the loading vector (i.e. wisc.pr\$rotation[,1]) for the feature concave.points_mean?

wisc.pr\$rotation ["concave.points_mean",1]

[1] -0.2608538

Q10. What is the minimum number of principal components required to explain 80% of the variance of the data?

```
tbl <-summary(wisc.pr)
which(tbl$importance [3,] > 0.8) [1]
```

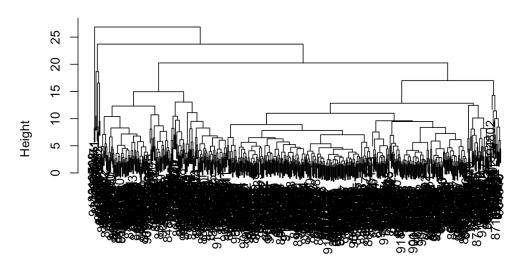
PC5 5

#Hierarichal clustering

The main function for Hierarchical clustering is called 'hclust()' it takes a distance matrix as input.

```
d <- dist(scale(wisc.data))
wisc.hclust <- hclust(d)
plot(wisc.hclust)</pre>
```

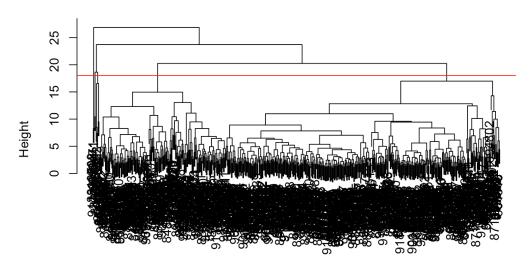
Cluster Dendrogram



d hclust (*, "complete")

```
plot(wisc.hclust)
abline(h=18, col="red")
```

Cluster Dendrogram



d hclust (*, "complete")

come back here. later to see how our cluster grps correspond to M or B groups

5. Combining methods

Her we will perform clustering on our PCA results rather than the original data.

In other words we will cluster using 'wisc.pr\$x' - our new better variables or PCs. We can chose as many or as few PCs to use as we like. It is your call!

```
wisc.pr$x [,1:3]
```

PC1 PC2 PC3 842302 -9.18475521 -1.946870030 -1.122178766

842517	-2.38570263	3.764859063	-0.528827374
84300903	-5.72885549	1.074228589	-0.551262540
84348301	-7.11669126	-10.266555635	-3.229947535
84358402	-3.93184247	1.946358977	1.388544953
843786	-2.37815462	-3.946456430	-2.932296681
844359	-2.23691506	2.687666414	-1.638471247
84458202	-2.14141428	-2.338186649	-0.871180715
844981	-3.17213315	-3.388831138	-3.117243065
84501001	-6.34616284	-7.720380945	-4.338098744
845636	0.80970132	2.656937673	-0.488400053
84610002	-2.64876984	-0.066509405	-1.525113359
846226	-8.17783882	-2.698602007	5.725193241
846381	-0.34182514	0.967428026	1.715661982
84667401	-4.33856172	-4.856809832	-2.813639832
84799002	-4.07207318	-2.974443983	-3.122526655
848406	-0.22985277	1.563382114	-0.801813642
84862001	-4.41412695	-1.417423150	-2.268323145
849014	-4.94435304	4.110716528	-0.314472408
8510426	1.23597583	0.188049490	-0.592761930
8510653	1.57677384	-0.572304625	-1.799863016
8510824	3.55420904	-1.661487969	0.450790799
8511133	-4.72904972	-3.302058265	-1.465247427
851509	-4.20482441	5.123858059	-0.751740573
852552	-4.94528075	1.542395143	-1.711687827
852631	-7.09232236	-2.016835734	-0.028984132
852763	-3.50717666	-2.169715994	-3.891122637
852781	-3.06136021	1.874902631	2.579478032
852973	-4.00374127	-0.536769860	-2.759197831
853201	-1.71380176	1.522365502	0.146058578
853401	-6.05411853	0.756511800	-0.348932376
853612	-2.89968469	-4.001774373	-2.998823809
85382601	-4.55077848	-0.337239419	-0.753637681
854002	-4.98621538	1.131593223	-2.575118337
854039	-2.98271631	-0.757756497	-3.490690279
854253	-2.76393718	0.354044421	-1.895295337
854268	-1.29505925	-0.912393466	-1.574383382
854941	3.74601730	1.412230504	1.730784885
855133	0.99719148	3.348346731	4.301788720
855138	-0.76459136	-0.885464837	-2.702540362
855167	2.14906252	1.922300194	-1.179446370
855563	0.09324934	-2.258764532	-1.903051997
855625	-9.08001023	-2.016898441	2.139343894
856106	-0.98958304	-0.984064148	-2.308702241
85638502	0.29328849	0.136978564	-3.299029123
857010	-5.37620991	0.134758404	-1.677366657

http://localhost:7017/ Page 18 of 33

85713702	4.57790860	-1.482915508	0.042037623
85715	-1.69851237	-2.350203862	-3.075383081
857155	2.13456708	-0.095745364	-1.491099878
857156	1.56610099	1.207370855	-0.368834570
857343	3.53979091	1.281368110	1.017315805
857373	3.15503795	1.687473796	-0.495857416
857374	3.44745515	0.497780720	-0.152383875
857392	-3.29964761	1.129943574	0.825666100
857438	0.67402621	2.114549137	-1.394992302
85759902	2.85564577	-0.152588874	0.428557855
857637	-4.64465207	2.308301508	-0.728168145
857793	-2.17494916	-0.971261438	-1.942179421
857810	3.71818738	1.786070097	1.278807604
858477	4.13232693	-2.401679261	-1.241865694
858970	2.38373822	-2.755233814	1.213018303
858981	2.57661610	-3.135912650	1.399049169
858986	-4.75492832	-3.009032875	-0.166094576
859196	2.31209785	-3.265116721	2.793766025
85922302	-1.69012080	-1.539322116	-1.798262954
859283	-1.81071217	-0.722104815	-1.464960560
859464	2.78347559	-2.308617372	0.472309720
859465	3.51555502	0.657730736	0.590855071
859471	-4.32619605	-9.194435552	1.491650337
859487	3.25841241	0.937013647	0.204949679
859575	-2.70221851	4.433240985	0.307075192
859711	0.30758513	-7.381317396	3.815728542
859717	-5.49886689	-0.937500513	-2.208188675
859983	0.36139121	-0.119633821	-2.039161544
8610175	2.62766457	0.696696349	-0.583428281
8610404	-1.42691206	1.965372092	1.110961327
8610629	0.83378424	-1.963876860	0.789737653
8610637	-6.22541880	-0.919260690	0.089748341
8610862	-11.65845644	-4.744442593	2.004115633
8610908	2.01980045	0.254675746	-0.647302452
861103	1.63694460	-1.714440587	0.433824539
8611161	-1.16643527	-2.512305286	-1.904043397
8611555	-10.75977535	2.255997858	0.038801648
8611792	-5.03038488	-0.773728357	2.679979922
8612080	2.17255269	-0.496441169	-0.948257444
8612399	-3.28534463	1.666770480	0.146802675
86135501	-0.60707378	-0.162071919	1.635935419
86135502	-3.58041324	2.204721916	-1.703037100
861597	0.93333687	-0.926885919	-0.040128345
861598	-1.25849730	-1.014684255	-0.080521995
861648	1.58686770	1.618232954	0.309726682

http://localhost:7017/ Page 19 of 33

861799	-0.25227559	0.530884425	0.588809736
861853	2.84492835	2.891103732	0.527524677
862009	1.96322268	0.964308498	-0.081136715
862028	-2.77342525	-0.557509977	-0.605238702
86208	-4.39236314	2.121640892	0.579000549
86211	2.58980407	-0.213446120	1.034701089
862261	3.90090576	-1.189020421	1.550616806
862485	2.81575365	-0.367560567	-1.350247556
862548	-0.61573908	-0.638349988	-0.199743121
862717	0.43247968	1.390820523	-0.082646762
862722	4.55102442	-3.525683788	-0.215074716
862965	3.44515217	1.423370040	-0.088646005
862980	2.11410008	-1.847747782	-0.556184595
862989	2.68898618	-1.418813334	0.968298772
863030	-3.21109134	-4.043198634	-2.712997305
863031	0.74861266	-1.796058175	-0.567866143
863270	3.15622889	1.034838878	-1.489373670
86355	-13.04464395	-0.980650357	0.598286247
864018	2.22672053	-0.666823902	-1.090600727
864033	2.49063954	-2.596401708	1.668257671
86408	0.10325957	-2.278139015	2.107708229
86409	-3.58813694	-3.922881433	2.045207163
864292	1.34776279	-3.553098748	1.730668471
864496	2.50791098	-3.248461200	-1.686140767
864685	2.04423086	-0.304616836	-0.533369397
864726	2.09522559	-3.663872062	3.629007335
864729	-3.10779316	-1.568009873	-2.611638159
864877	-4.95236801	-2.382749650	-1.933405900
865128	-0.85026612	2.304707468	0.566963050
865137	2.96339173	-0.371179883	-2.064716427
86517	-3.33120209	1.324392029	0.727373652
865423	-12.88327621	-2.314585873	6.323133317
865432	0.77006610	0.064052860	-1.133411514
865468	2.20057925	0.734958544	0.266845303
86561	3.14064881	1.875758557	-0.669779705
866083	0.63831885	0.910564828	-2.140840449
866203	-1.91744627	3.534971966	1.253821297
866458	-1.40762980	-1.303782215	0.640864876
866674	-4.63960895	1.480714297	-1.169651383
866714	1.87581684	-1.421979452	-1.069870821
8670	-1.43081046	1.048681220	-1.260901861
86730502	-1.35143790	1.153126842	-1.208729712
867387	0.70849143	1.566853657	-0.669405084
867739	-2.16950708	2.823776101	-0.561613332
868202	1.97510247	0.419018755	-0.380240597

http://localhost:7017/ Page 20 of 33

868223	2.59850310	0.481911436	1.870726668
868682	3.25514318	0.417953651	-0.691197447
868826	-3.77870160	-0.859625218	3.667773435
868871	1.99028348	-1.328186508	1.134246527
868999	5.01059989	-0.574194748	-0.841697543
869104	-1.34383800	1.273650278	0.555354894
869218	2.45380972	-0.897665441	0.352823033
869224	1.83566648	0.090946742	-1.444010288
869254	4.34266862	0.892786708	0.638694049
869476	0.73216868	-3.698927716	0.657926113
869691	-2.39788795	-4.833735642	-1.027253164
86973701	-0.39275308	-1.082115915	2.245629588
86973702	0.41196533	0.389288784	-1.044851048
869931	3.04724369	2.235817239	-0.469040276
871001501	1.44129554	-0.305601552	0.551062974
871001502	-0.08311297	-7.144073768	-0.059957215
8710441	-7.08707084	-12.562140869	7.352238887
87106	3.74011364	-0.250281306	-0.097297641
8711002	0.96832079	-0.944113582	-0.890028129
8711003	2.41659356	-0.005547532	-0.797678511
8711202	-4.09718264	0.378470806	1.445046044
8711216	0.75094226	3.067949187	1.447907672
871122	3.65143365	0.674055644	-0.906188637
871149	4.67609710	1.102886823	-0.257081973
8711561	0.59725624	-1.784081072	1.480433343
8711803	-3.38435289	2.908477953	3.004594027
871201	-6.14447970	2.015878994	-1.564766888
8712064	1.32386648	-1.468002337	0.847675513
8712289	-5.48932277	4.162167037	-0.600732185
8712291	2.99476813	2.736453261	-0.161042080
87127	4.38287625		-0.947274735
8712729	-1.21256768	2.037246395	0.979834026
8712766	-5.06520424		1.172615291
8712853	1.97596280		-0.277115927
87139402	2.51266554		-0.496186180
87163	0.94665494		-0.350877607
87164	-2.81688978		-2.151251366
871641	3.48092345	-1.618268077	2.673117328
871642	4.65463399	0.222718846	1.557987797
872113	5.34691339	-1.025855214	
872608	-1.16986842	-7.008319996	4.268106508
87281702	-2.95370292		-1.041419103
873357	4.97132782	3.383227983	0.439451738
873586	4.06045289	1.245070643	0.573972801
873592	-9.50430706	5.598558052	-0.636679193

http://localhost:7017/ Page 21 of 33

873593	-8.99924714	_0 580520281	-3.131959404
873701	-0.75821134	1.607118481	-1.513763762
873843	2.65702172	-0.539461132	1.705514033
873885	0.38967611	0.988372689	-2.582896564
874158	3.88564657	-0.815354285	1.014824328
874217	-0.36455538	3.571318842	-2.222023387
874373	2.83339783	0.398379162	-0.347139642
874662	3.30737348	-0.155604280	0.541002917
	3.35435293	1.102762109	0.151956099
874839			-0.326136731
874858	-6.51738201	-8.004126824	
875093	1.71622508	0.542688019	3.295475582
875099	5.56084294	0.477427493	4.123834207
875263	-1.77809695	-2.774146213	-0.941934046
87556202	-2.60918908	-1.560049030	-0.217995141
875878	2.81656023	0.969257722	-0.382255094
875938	-2.49624245	-2.276479790	1.318318594
877159	-1.27590451	2.441111044	2.464148795
877486	-3.47014398	2.275846405	-1.237595112
877500	-1.25557034	-0.382057094	-3.834946034
877501	1.47213604	-0.116787496	-0.214479145
877989	-1.64801062	2.100442902	-0.974414210
878796	-9.02864522	0.654596849	0.221085545
87880	-4.55058603	-3.083925626	-2.268408346
87930	0.78050321	-0.652275325	-0.643091145
879523	0.22289648	0.701204504	-2.268765765
879804	3.45185811	-1.305789765	-0.029238791
879830	-0.44615183	2.785257008	0.443991604
8810158	-0.31416180	-2.075734301	-2.081480386
8810436	2.05738166	2.470613908	-1.460433790
881046502	-4.80474158	3.026440019	2.741606351
8810528	2.99607430	0.396429212	-0.595941573
8810703	-12.27421974	7.536778599	10.103533703
881094802	-3.36923202	-2.585550416	7.085840019
8810955	-2.50656080	-2.612349885	0.206198902
8810987	-1.31690973	-2.152585453	-1.785061911
8811523	0.38841453	-2.274796314	0.626780574
8811779	2.75450931	-1.085879226	1.705737291
8811842	-4.93923501	2.845820436	0.913075404
88119002	-4.29429880	4.662172958	-0.006834994
8812816	2.37141521	0.732757814	-1.524685925
8812818	0.96634003	-0.438052741	-1.446423567
8812844	2.97318270	-1.809381711	-0.223238842
8812877	-1.80166175	-0.166314618	-2.626288708
8813129	2.47618073	1.417329346	-0.316254972
88143502	0.82531310	1.249148354	0.267725157

http://localhost:7017/ Page 22 of 33

88147101	3.89127342	-0.538097276	-0.880783727
88147102	0.38923560	0.613875324	-1.351679911
88147202	1.45604019	0.201549415	-1.097912687
881861	-2.75314552	-3.462726769	-2.471875866
881972	-3.25439115	0.125090426	-2.893551435
88199202	4.01560693	1.353056834	-0.360155815
88203002	3.67318128	1.290493562	0.424237023
88206102	-3.31340866	3.935689268	-0.205273062
882488	4.03305870	-1.161524089	0.189581014
88249602	2.33218128	1.347793816	0.232482459
88299702	-8.39645911	4.150252991	0.046927784
883263	-2.64524110	3.947699392	1.370675284
883270	0.68116951	1.134992457	1.232581929
88330202	-3.34735048	2.152675899	-1.107968825
88350402	2.22576208	1.213931555	-0.192244621
883539	4.47364832	1.739598699	-0.887426246
883852	-1.32145867	-4.785266054	-0.661952674
88411702	1.84846341	1.582735509	1.262806090
884180	-3.79222888	1.025330459	0.897845626
884437	2.27931254	-2.075859442	2.428512195
884448	3.22593164	1.171043084	0.264431658
884626	-1.02499890	-2.359186109	-1.615464155
88466802	2.58028546	-0.728572047	-0.547660903
884689	2.48369783	-0.460191483	-0.598393817
884948	-7.13756174	2.073018195	1.087430513
88518501	3.02981522	0.648897838	0.310409540
885429	-7.07878207	-0.527758800	-1.278861780
8860702	-1.31294375	1.773865017	-1.671161321
886226	-3.70838653	2.805008998	-1.754649703
886452	-0.46007427	-0.393819769	-1.003815294
88649001	-6.38703001	1.821491443	0.479427881
886776	-5.25092469	-3.891107226	-0.833120389
887181	-8.72618338	-3.276993966	2.901146761
88725602	-3.68807276	-1.064709236	-3.512498419
887549	-3.37528098	3.368264175	-1.765224543
888264	0.83947101	3.497008034	0.182323228
888570	-3.28148526	0.989802279	2.552140777
889403	1.90709089	3.119201237	-1.062121558
889719	-1.61336234	2.472397898	-1.989384336
88995002	-6.61390681	5.997998811	0.988507976
8910251	1.46338831	-1.685149061	1.168877382
8910499	2.01038350	1.102516460	1.336847795
8910506	2.28233204	-0.009485293	-0.450043551
8910720	0.73909276	-3.149749158	1.468565664
8910721	4.50260070	3.166486414	-0.187896260

http://localhost:7017/ Page 23 of 33

8910748	3.15597737	-0.409941090	-1.107842420
8910988	-7.66940121	3.072602737	1.480656591
8910996	3.63593673	-1.588189625	0.715650110
8911163	-0.33704893	3.141609173	-0.115619749
8911164	1.27546111	-0.848306563	3.067369195
8911230	4.34219338	0.321694999	0.388133185
8911670	-0.01741547	3.456302110	0.306650981
8911800	3.48935577	2.631766723	-0.072124181
8911834	1.86490793	0.901271186	-0.870402828
8912049	-4.95590744	1.339944038	-1.455772409
8912055	2.46855062	0.137904923	1.793307382
89122	-3.78173258	1.900235409	-1.742708503
8912280	-2.63493731	-0.576647607	-1.992946242
8912284	1.37147827	-0.005800358	1.547927203
8912521	4.13524399	1.375298123	0.877276691
8912909	0.68633639	-1.693519506	1.025310679
8913	4.30833870	1.976746804	-0.251085558
8913049	-0.40295486	-3.720160654	4.582436940
89143601	3.00096526	-0.353848878	1.232930128
89143602	-3.13125303	-4.269702670	5.861122345
8915	0.41746954	0.807865734	-1.092330348
891670	1.41897459	-1.392752638	-0.836251021
891703	2.86806475	0.268408338	-0.468574962
891716	3.36288672	0.806890637	-0.806369895
891923	3.31208587	1.440908382	-0.583119519
891936	4.76551701	0.542484530	0.893082182
892189	2.38378528	0.823391074	1.982219343
892214	2.78464581	2.530807159	-0.883257225
892399	3.39085817	-0.753301131	1.770235939
892438	-6.58444233	1.483557076	0.361022239
892604	1.45405111	-0.591422478	1.145332613
89263202	-7.17661843	-0.055097580	2.011148288
892657	3.57568917	-0.890799182	0.261257526
89296	3.05434514	0.179055836	0.621226832
893061	3.07173303	0.305789373	1.258146621
89344	3.85015201	1.523386003	-1.143503482
89346	5.38551771	-0.555766489	0.559577694
893526	4.65358767	3.062692587	-0.075942289
893548	4.02400169	2.540671520	0.352906609
893783	3.34487880	0.068502773	-0.269955756
89382601	3.29694105	3.135950089	0.148761476
89382602	2.13452598	0.004388987	-0.055090483
893988	4.03446724	0.240486452	-1.326959165
894047	3.76210430	-4.394333943	3.946735131
894089	4.86695697	2.337158334	-0.168146877

894090	4.74307511	1.796820963	0.169570993
894326	-1.89365006	2.390134936	-1.675017462
894329	-1.57154834	-6.503377021	2.165253786
894335	4.02308037	1.400973753	3.183698584
894604	1.29372783	-3.467757378	1.108552569
894618	-1.83322742	4.317723823	0.316406925
894855	1.80878902	-0.395858988	-1.229252510
895100	-7.23012357	0.035670411	-2.838568987
89511501	3.14288478	0.741873771	-0.861593396
89511502	2.88472980	0.464323069	-0.475325102
89524	3.14688506	1.769691255	-0.035802185
895299	4.62431058	1.834004631	0.064174505
8953902	-2.24853190	0.348229535	-1.063814825
895633	-2.10594822	-1.120987593	1.773597643
896839	-1.78437991	0.268984957	-1.721001037
896864	0.52723271	-1.264758769	-0.919631937
897132	3.21067592	-1.102473082	1.610564041
897137	4.38256677	0.760434627	-0.400245248
897374	3.80666840	0.909118626	0.442778081
89742801	-3.31675558	1.575666422	1.359532256
897604	2.39117199	-0.989160421	-0.756844005
897630	-4.67605183	0.967741443	-0.291016193
897880	3.06499827	-1.134839030	0.374293207
89812	-7.34675421	5.238014242	0.316677175
89813	-0.29505357	-0.226735533	-1.088458865
898143	1.93726491	-2.542750760	-1.000543159
89827	2.07535283	-1.804931340	-1.166053416
898431	-4.97078859	1.331400687	0.870211875
89864002	2.17500311	-0.958068260	-0.234058923
898677	2.42941383	-3.444173357	3.453905770
898678	3.36500493	0.562434015	-0.192404359
89869	1.20553704	1.315886497	-0.929922572
898690	3.32315522	-0.474753848	-0.897496041
899147	1.92709976	-1.460184989	-0.100554033
899187	4.03828183	1.355724562	0.537868401
899667	-6.17684877	-5.103514894	0.520682438
899987	-10.92468559	3.699998472	-0.890910165
9010018	-2.11663452	-0.296368297	0.866722769
901011	3.00954689	-0.243545681	1.981989096
9010258	0.98178471	-0.796551513	2.367338148
9010259	-0.35757923	-2.125968004	0.707586438
901028	3.19111003	1.847526445	0.099310968
9010333	2.99060189	-1.629958015	2.554470244
901034301	3.19110631	-0.578830133	0.489333658
901034302	4.58792496	2.758699732	0.979680613
-			

http://localhost:7017/ Page 25 of 33

001041	2 40272672	1 107004102	0 240470172
901041 9010598	2.48373073 2.39655663	1.187894183 0.250524355	0.348470172 -0.326850211
9010398	0.77892982	2.122475872	0.320530211
9010872	3.13722800	1.486714010	-1.197824530
9010877	-2.73260737	3.941699950	-0.017154393
901000	-6.22155168	1.388887616	2.731833465
9011494	2.22182078	0.356846723	-1.100713817
9011495	-5.30247440	6.717504288	1.484358772
9011971	-7.24162919	3.652255530	-0.155565119
9012000	-4.20352196	-1.175208822	-0.155505119 -1.588620367
9012515	2.49743889	2.016761090	-0.730631718
9012306	-3.63181933	1.954720962	-0.730031718 -0.828046542
9012795	-3.03161933 -3.51615443	3.855214937	-0.020040342 -1.021630135
		0.000==.007	-1.021030133
9013005	2.61264945		
901303	0.09602363		-1.669994689
901315	-2.04305926	-6.421169887	1.273001142
9013579	3.06405348		-0.278817341
9013594	1.60217670	-0.292338995	-1.584830948
9013838	-3.76262144	-5.980033549	-4.876070721
901549	0.92475292	-2.300457270	-1.594708683
901836	3.28545874	-0.201134359	-0.627204410
90250	1.55348525	-0.978996316	1.195507876
90251	0.39842934	-2.159312724	-0.944773717
902727	2.29854534	0.931258175	-1.128959899
90291	0.54242409	1.315910048	0.994879134
902975	2.05700950	-0.320153445	-0.122350756
902976	3.42150749	2.443368505	0.045009146
903011	0.47199125	-3.699574203	1.952067274
90312	-4.69936579	0.195783486	3.753730061
90317302	3.64017727		-1.161124120
903483	3.59152918		1.850401478
903507	-4.14544978		-0.887194062
903516	-7.66539681		-2.151756805
903554	2.00861205		-0.103997306
903811	2.91826573		0.295467269
90401601	0.50029759	-0.106637635	-1.027843564
90401602	1.90981814	0.650821226	1.449013512
904302	3.78177167	0.325859645	-1.057150017
904357	2.70601296	0.219472629	-0.314228262
90439701	-6.28414409	-2.034757025	-3.179284831
904647	3.42775463		-0.731051348
904689	1.99962900	0.293068594	0.618659534
9047	2.47111877	0.334752480	-1.481058853
904969	3.64405137	1.240273815	1.385405524
904971	2.34412427	-0.682141105	1.131738080

905189	1.14199209		-1.239628958
905190	1.34966974	-0.369249002	2.484489568
90524101	-2.76094145		-1.178586912
905501	1.96759233	0.175693916	0.420361503
905502	2.92861019	0.494421191	-1.016176763
905520	2.92380175	-0.377000536	-0.886014238
905539	3.99267514	-0.959012260	0.204224433
905557	-0.15500167	0.437623166	-0.461040030
905680	0.93326963	2.104094210	1.431657320
905686	1.92792536	-0.891682104	0.042307522
905978	2.62003325	-2.499901968	2.432858840
90602302	-5.99833869	0.090949196	1.709697317
906024	2.91930912	0.009204582	-0.900511128
906290	3.33900683	0.022205192	0.502858739
906539	2.01314368	-0.777767778	0.125468395
906564	-1.98671401	-2.314011967	1.235217081
906616	1.66813615	-0.861560770	-1.009417991
906878	0.37085180	-0.113512779	-0.997485893
907145	1.84733447	-2.538995601	2.003584668
907367	4.69890766	0.431922385	0.005300424
907409	1.28205907	-2.548666752	-0.083786622
90745	2.31221232	-0.401678043	0.274381769
90769601	4.65498928	0.781611021	0.191953023
90769602	3.94000659	2.028231657	0.691127602
907914	-4.94550694	-3.003421435	-1.752529645
907915	0.91499842	-2.476833776	0.362287911
908194	-4.54511193	0.815283142	-0.518299038
908445	-4.43843746	0.991566424	1.441051113
908469	2.19305464	1.803766114	-0.434726892
908489	-0.66385679	-0.436476802	-2.678425146
908916	2.23890605	0.454189872	-0.465118333
909220	2.12331215	1.193746541	0.489940891
909231	2.61301823	1.830449830	0.763657618
909410	3.20814515	2.233177743	-0.166078298
909411	0.30648073	-2.183077955	0.039251705
909445	-2.47018342	1.498506224	0.590933256
90944601	3.45168126	2.134583340	0.196989342
909777	3.89709781	-0.730133530	2.134193239
9110127	-0.98086448	2.208486712	-1.698424795
9110720	1.25614629	-1.066752885	0.259247160
9110732	-3.24256054	1.776795240	-1.791577273
9110944	1.55585043	1.037014643	-1.300101127
911150	1.24815985	1.587442008	-0.326126484
911157302	-4.32784530	4.045772565	-0.192336290
9111596	1.07650412	-1.802032564	1.779475879

http://localhost:7017/ Page 27 of 33

9111805	-2.50954497	2.526584455	0.797500590
9111843	2.21514321	-0.029865010	0.635900610
911201	1,17279871	0.474422064	-0.481147553
911202	2.83380799	1.017230750	-1.485323075
9112085	1.85545330	1.570011674	0.825756166
9112366	1.32744636	-0.776412499	1.255653387
9112367	2.80030738	1.664836753	-0.193847761
9112594	3.17177797	2.073908798	0.074834422
9112712	4.08348433	0.484283961	0.630808963
911296201	-3.47243837	1.671413723	0.856073545
911296202	-16.30488665	7.769016888	6.229933870
9113156	2.55786437	2.491852880	0.170162861
911320501	2.96627878	0.068786611	-0.777727248
911320502	2.75500634	1.792341696	-0.757147909
9113239	-1.36913518	-2.108156394	-0.852969723
9113455	0.41729731	-0.116417262	-0.535671262
9113514	3.83502421	-0.898388162	0.699948069
9113538	-5.92575082	-1.227216861	1.720074272
911366	-0.64491119	-3.422999542	0.607750202
9113778	2.68148698	-1.442286469	0.339105123
9113816	2.03986918	0.902461858	2.840741857
911384	1.39806620	1.770668244	-1.478260847
9113846	3.53330636	1.245758018	3.382434941
911391	1.98877953	-1.897792693	-2.029212586
911408	1.99642872	0.206189711	-1.639882150
911654	0.52020039	0.972512103	-0.323850426
911673	3.16958159	2.087215768	-0.977945244
911685	2.20077374	-1.284908734	-0.886619084
911916	-3.82107090	-2.303209735	0.027521977
912193	2.98286009	0.672747298	-0.589970772
91227	2.47887396	2.361880555	-0.432234246
912519	1.27099312	-0.509385643	-2.237023334
912558	2.23319191		-0.900644424
912600	0.05613371		-2.253532782
913063	-2.31535334		1.776847691
913102	2.28232306		-0.820466562
913505	-4 . 75043499		-1.252648906
913512	1.72479328		-0.080167877
913535	0.74112965		-1.995953713
91376701	2.89176621		-0.804935982
91376702	1.65306732	4.551659142	0.751661685
914062	-3.10079133	1.235066483	1.286223339
914101	4.06069761	0.560461075	2.174771231
914102	2.79087498	1.076940539	0.843799658
914333	1.47362330	1.589695437	0.319710133

http://localhost:7017/ Page 28 of 33

914366	0.22919297	-1.514570755	-2.214334625
914580	2.55101378	0.763228924	-1.607988258
914769	-3.71243855	1.057755506	0.521494703
91485	-5.09036286	2.017432706	-0.708746074
914862	0.60008810	0.837717560	-0.686437977
91504	-2.78794673	-3.382584454	-0.752414277
91505	1.40753273	-1.504446838	-0.353584900
915143	-7.25279714	5.490904302	1.909245947
915186	-1.29772469	-7.723057325	-0.688309017
915276	-1.07574673	-8.287483337	0.433810751
91544001	1.24866183	-1.594281055	-1.355282076
91544002	1.25582341	-4.113567176	-0.412366931
915452	1.16550969	1.664229389	-0.737229323
915460	-4.09054331	-2.800463137	-0.506987325
91550	1.88597978	-1.670721680	0.119692299
915664	2.76500279	2.159148972	-0.021077623
915691	-2.22503669	-1.939927827	-2.283594176
915940	1.13058210	1.409761524	-0.564556376
91594602	0.73312738	1.941988438	-0.274994234
916221	2.33161916	-0.789451973	-0.349735677
916799	-2.69485257	1.942001189	-0.959719996
916838	-3.37604410	2.331377578	-1.200875536
917062	0.19969487	-1.075413752	-0.672557052
917080	1.17499204	-1.010481061	-0.711142576
917092	1.29150016	-4.959860548	0.345622908
91762702	-8.62316838	3.456410638	-0.178899380
91789	4.42556234	0.785345015	-0.473002900
917896	0.61929383	-0.635789842	-0.488421490
917897	3.24945659	-1.284794677	-0.218284421
91805	3.34997942	-2.670950443	-1.777523046
91813701	0.85904712	-0.096763381	-2.811570840
91813702	3.15384692	0.870776432	-2.193784489
918192	-0.34622347	-1.539870352	2.843951628
918465	2.45924278	-0.600319751	-1.170435277
91858	1.52875558	-0.404861554	0.603933344
91903901	1.77202587	-0.803503014	-2.026135257
91903902	2.67808519	1.483403662	-1.740162222
91930402	-4.02489378	2.938844225	1.658026507
919537	2.13447312	-1.517245623	0.178807901
919555	-5.16086993	2.380108987	-0.421906048
91979701	-0.53546080	-0.380380451	-0.442745212
919812	-0.34282174	-3.531373815	0.069607554
921092	4.19339024	-2.365311059	1.325978010
921362	1.14182718	-5.594535868	1.299893249
921385	1.66401100	-2.387517361	1.500928782

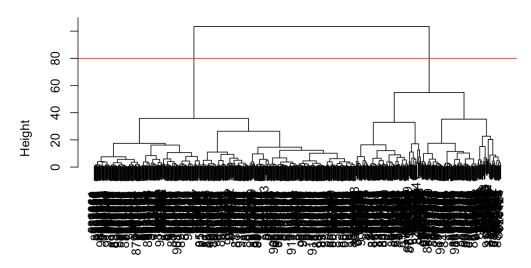
http://localhost:7017/ Page 29 of 33

```
921386
           -1.01082308
                        -1.091429307 -0.632142163
921644
            1.29978604
                         1.819814057 0.372979316
            2.37134219
922296
                         1.680097929
                                      0.384190094
922297
            1.66440651
                         0.213774641 - 0.147942249
922576
            1.92598353
                          1.136739705
                                       0.477781900
922577
            4.23349159
                        -0.184110499 -0.326131445
922840
            2.67551655
                        -2.313756961 -0.053800867
923169
            3.83312511
                        -0.495813665
                                      0.922428047
923465
            2.54919727
                        -0.228129228
                                       1.412934778
923748
            4.69079604
                         0.766803238
                                       1.542607509
923780
            2.02325691
                        -1.260133116
                                       0.504482203
924084
            2.89340232
                         1.450359601
                                       0.779859803
924342
            3.49912218
                        -1.799249342
                                       2.764024589
924632
                         0.829339088
                                       0.564300199
            2.15201013
            2.05327740
924934
                        -1.615038205
                                       1.837342797
            3.87388097
924964
                        -1.083301553
                                       1.858308794
925236
            4.06028949
                        -0.122061034
                                       3.235925374
925277
            0.09858059
                         0.213372093
                                       0.388587548
925291
            1.08841850
                        -1.291711328
                                       1.428122289
925292
                         0.177863190
            0.48134743
                                       1.031200235
925311
            4.86602793
                         2.129232607
                                       3.411187296
925622
                        -3.479575000 -3.259923297
           -5.91241029
926125
           -8.73365338
                         0.573350185
                                      0.896301447
926424
           -6.43365455
                         3.573672989
                                       2.457324373
           -3.79004753
                         3.580897052
926682
                                       2.086640366
926954
           -1.25507494
                         1.900624364
                                       0.562235817
927241
          -10.36567336
                        -1.670540206 -1.875379194
92751
            5.47042990
                          0.670047220
                                      1.489132801
```

```
d.pc <- dist(wisc.pr$x[, 1:3])
wisc.pr.hclust <- hclust( d.pc, method="ward.D2")
plot(wisc.pr.hclust)
abline(h=80, col="red")</pre>
```

http://localhost:7017/ Page 30 of 33

Cluster Dendrogram



d.pc hclust (*, "ward.D2")

```
grps <-cutree (wisc.pr.hclust, h=80)
table(grps)</pre>
```

grps 1 2 203 366

We can use 'table() function to make a cross-table as well as just a count table

```
table(diagnosis)
```

diagnosis B M

357 212

```
table(grps, diagnosis)
```

```
diagnosis
grps B M
1 24 179
2 333 33
```

Write a note here about how to read this cross-table result. The results indiciate that our cluster 1 mostly captures cancer (M) and our cluster 2 mostly captures healthy (B) individuals.

##7. Prediction

and plot this up

#url <- "new samples.csv"</pre>

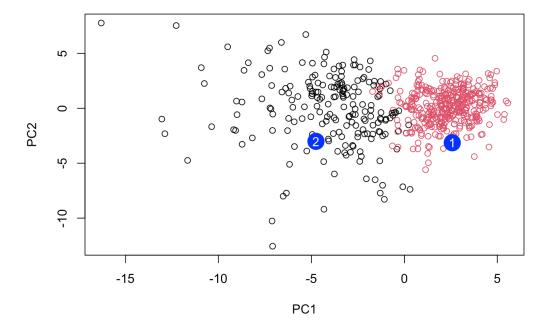
```
url <- "https://tinyurl.com/new-samples-CSV"</pre>
new <- read.csv(url)</pre>
npc <- predict(wisc.pr, newdata=new)</pre>
npc
           PC1
                     PC2
                                 PC3
                                            PC4
                                                      PC5
PC6
           PC7
[1,] 2.576616 -3.135913 1.3990492 -0.7631950 2.781648
-0.8150185 - 0.3959098
[2,] -4.754928 -3.009033 -0.1660946 -0.6052952 -1.140698
-1.2189945 0.8193031
            PC8
                      PC9
                                 PC10
                                           PC11
                                                     PC12
PC13
         PC14
[1,] -0.2307350 0.1029569 -0.9272861 0.3411457 0.375921
0.1610764 1.187882
[2,] -0.3307423 0.5281896 -0.4855301 0.7173233 -1.185917
0.5893856 0.303029
          PC15
                                  PC17
                                              PC18
                                                           PC19
                     PC16
PC20
[1,] 0.3216974 -0.1743616 -0.07875393 -0.11207028 -0.08802955
-0.2495216
[2,] 0.1299153 0.1448061 -0.40509706 0.06565549 0.25591230
-0.4289500
           PC21
                      PC22
                                  PC23
                                             PC24
                                                          PC25
PC26
     0.1228233 0.09358453 0.08347651 0.1223396 0.02124121
[1,]
[2,] -0.1224776 0.01732146 0.06316631 -0.2338618 -0.20755948
-0.009833238
             PC27
                         PC28
                                       PC29
                                                    PC30
[1,] 0.220199544 -0.02946023 -0.015620933 0.005269029
[2,] -0.001134152  0.09638361  0.002795349  -0.019015820
```

points(npc[,1], npc[,2], col="blue", pch=16, cex=3)

plot(wisc.pr\$x[,1:2], col=grps)

http://localhost:7017/
Page 32 of 33

text(npc[,1], npc[,2], c(1,2), col="white")



Q18. prioritize patient 2 because it shows the most malignant characteristics.

http://localhost:7017/ Page 33 of 33