

AlbuminLevel

Daniel Mejia

2022-06-29

Albumin Level

Analysis of Albumin level and Vitamin D and B12 deficiency. Looking at Age, Gender and BMI

The objective is to analyze how the Albumin and Vitamin D and Vitamin B12 are related.

Reading original data

For this document we are going to use the file with 1587 rows.

```
data <- read_excel("MayaMD_1587.xlsx")
```

Convert to data frame

After reading the excel document we need to convert to data frame.

```
df <- data.frame(data)
```

Add age group

We need an attribute with the age group for every row. This attribute is added to the data frame.

```
df$Age.Group <- as.integer( ( as.integer( format( Sys.Date(), "%Y" ) ) - as.integer(format(as.Date(df$D
```

Add BMI group

For the MBI plots we need to group the BMi values

```
df$BMI.Group <- as.integer(as.integer(df$BMI) / 10) * 10
```

Convert values to numeric all the required attributes

All the values come in string format, we need to convert to numeric values.

```
## Warning: NAs introduced por coerción
```

Ranges

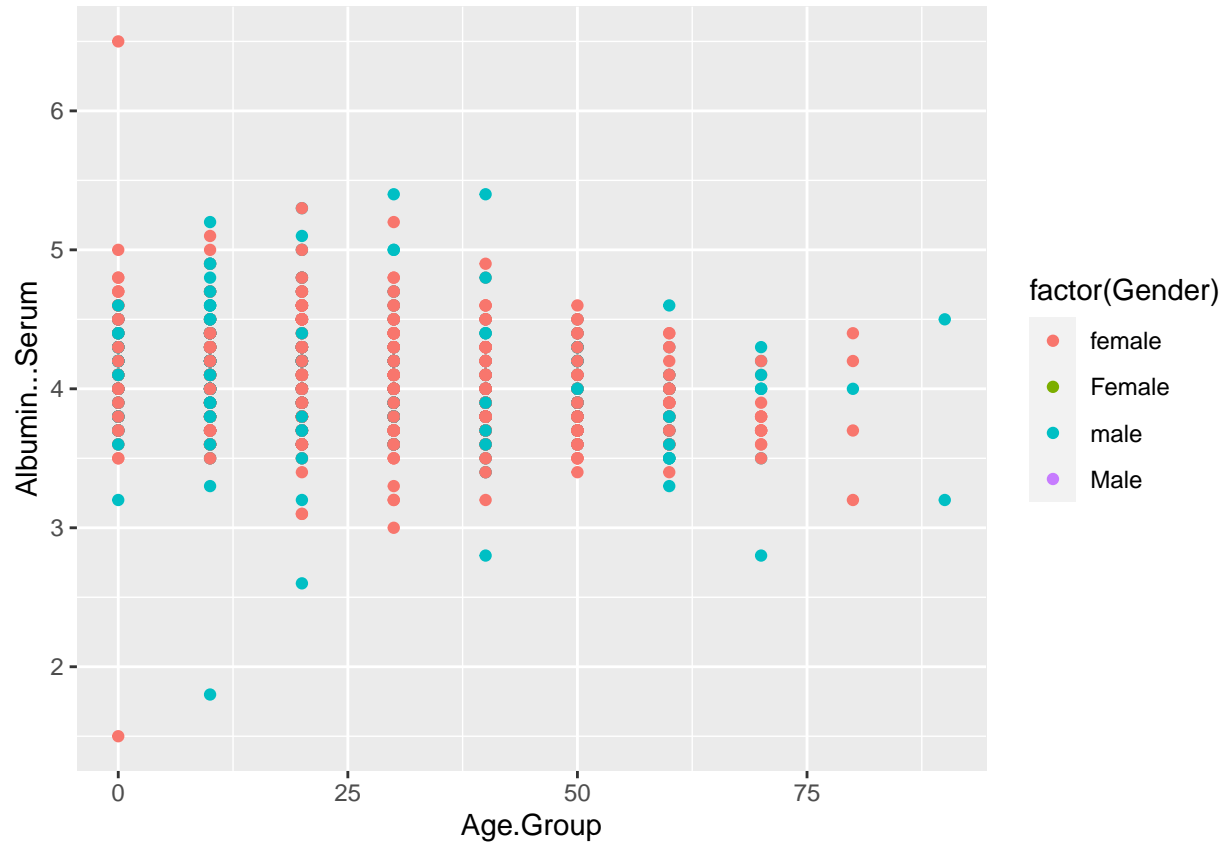
Albumin Serum < 3.5 g/dl low P. Diagnostic: Hypoalbuminemia

```
albumin_df <- df_mod %>% filter(!is.na(Albumin...Serum) & !is.na(Vit.D.assay) )
```

Albumin Vit.D Plots

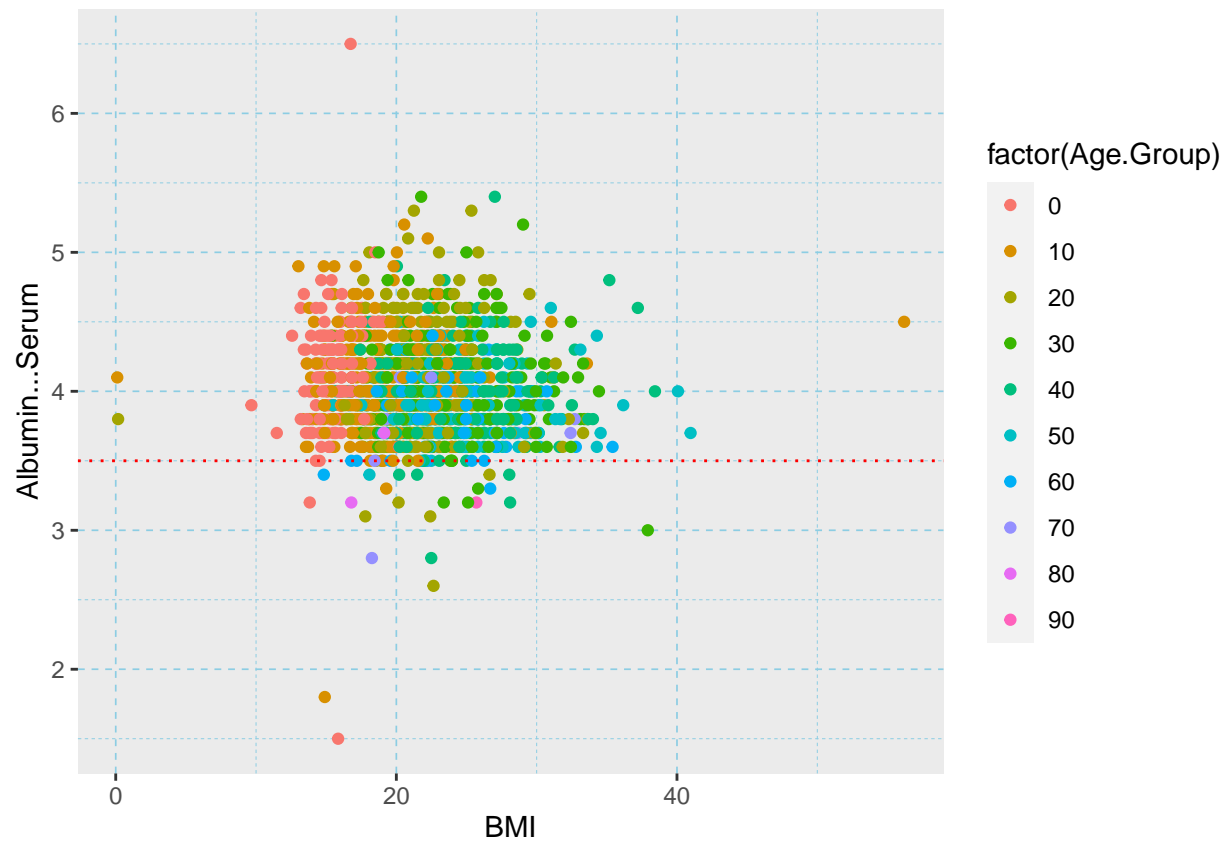
Albumin By Age Group

```
albumin_df %>% ggplot(aes(x=Age.Group, y=Albumin...Serum)) + geom_point(aes(colour=factor(Gender)))
```



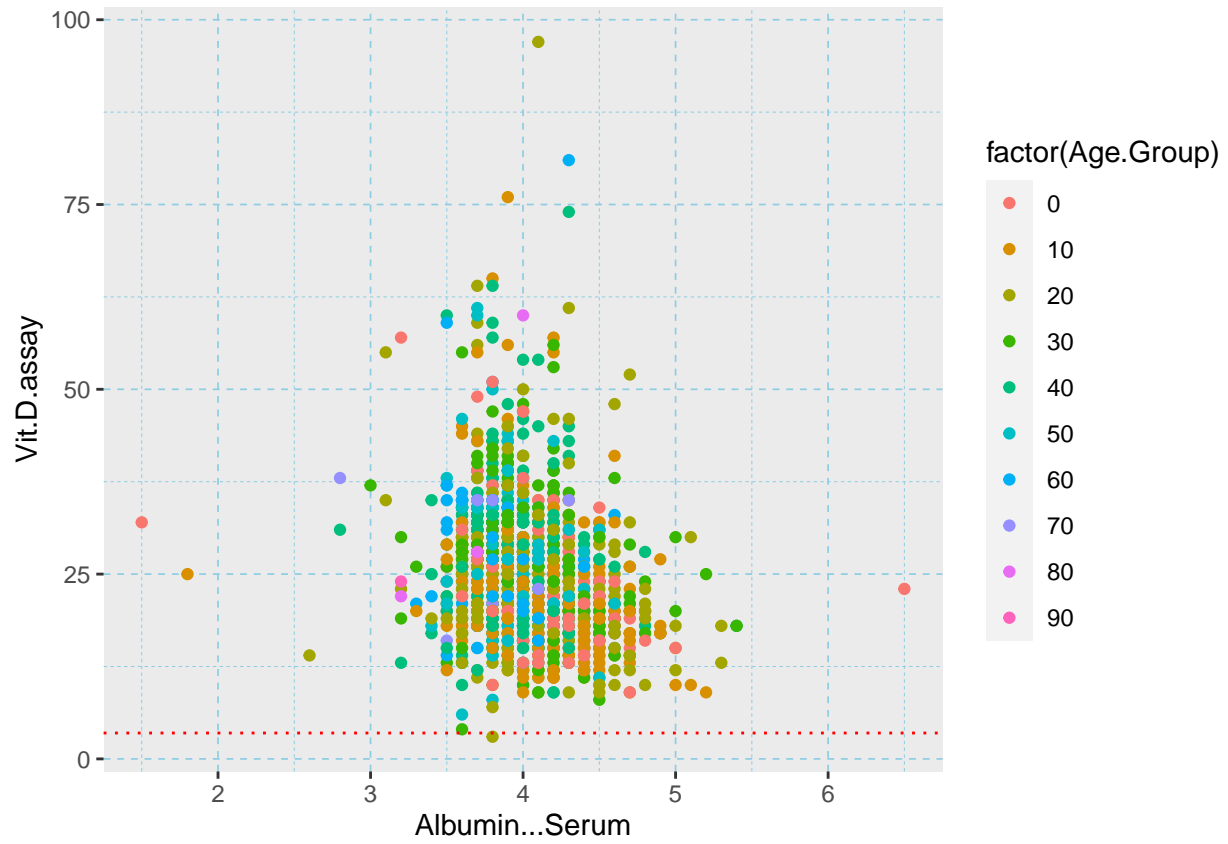
BMI - Albumin Plot

```
albumin_df %>% ggplot(aes(x=BMI, y=Albumin...Serum)) + geom_point(aes(colour=factor(Age.Group))) + theme_minimal()
```



Ablumin - Vit.D Plot

```
albumin_df %>% ggplot(aes(x=Albumin...Serum, y = Vit.D.assay)) + geom_point(aes(colour=factor(Age.Group)))
```



Detail data of people with Albumin < 3.5

```
albumin_df %>% filter(Albumin...Serum < 3.5 ) %>% select(Albumin...Serum, Vit.D.assay, Gender, Village
```

##	Albumin...Serum	Vit.D.assay	Gender	Village	Age.Group
## 1	3.2	22	female	Umkatieh A, B, C, D	80
## 2	1.5	32	female	UMROI MADAN	0
## 3	3.4	17	female	UMROI JAIWA PDENG	40
## 4	3.2	13	female	Mynsain	40
## 5	3.0	37	female	Umshaproh	30
## 6	3.1	55	female	Lumblei	20
## 7	3.4	22	female	Itsohpair	60
## 8	3.4	35	female	Itsohpair	40
## 9	3.1	35	female	Itsohpair	20
## 10	3.3	26	female	Itsohpair	30
## 11	3.2	30	female	Umden Arka	30
## 12	3.2	19	female	Umden Arka	30
## 13	3.4	18	female	Umket	50
## 14	3.4	19	female	Pynthor	20
## 15	3.2	23	male	Umkatieh A, B, C, D	20
## 16	1.8	25	male	UMROI MADAN	10
## 17	3.3	21	male	Lumdaitkhla	60
## 18	2.8	31	male	Nongtraw	40
## 19	3.2	57	male	Mynsain	0
## 20	3.3	20	male	Palwi	10
## 21	2.6	14	male	Palwi	20
## 22	3.4	25	male	Umshaproh	40
## 23	3.2	24	male	Itsohpair	90
## 24	2.8	38	male	Mawbri	70

Detail data of people with Vit.D < 3

```
albumin_df %>% filter(Vit.D.assay < 3.0 ) %>% select(Albumin...Serum, Vit.D.assay, Gender, Village, Age, Family.Id)

## [1] Albumin...Serum Vit.D.assay      Gender      Village
## [5] Age.Group      Family.Id
## <0 rows> (or 0-length row.names)
```

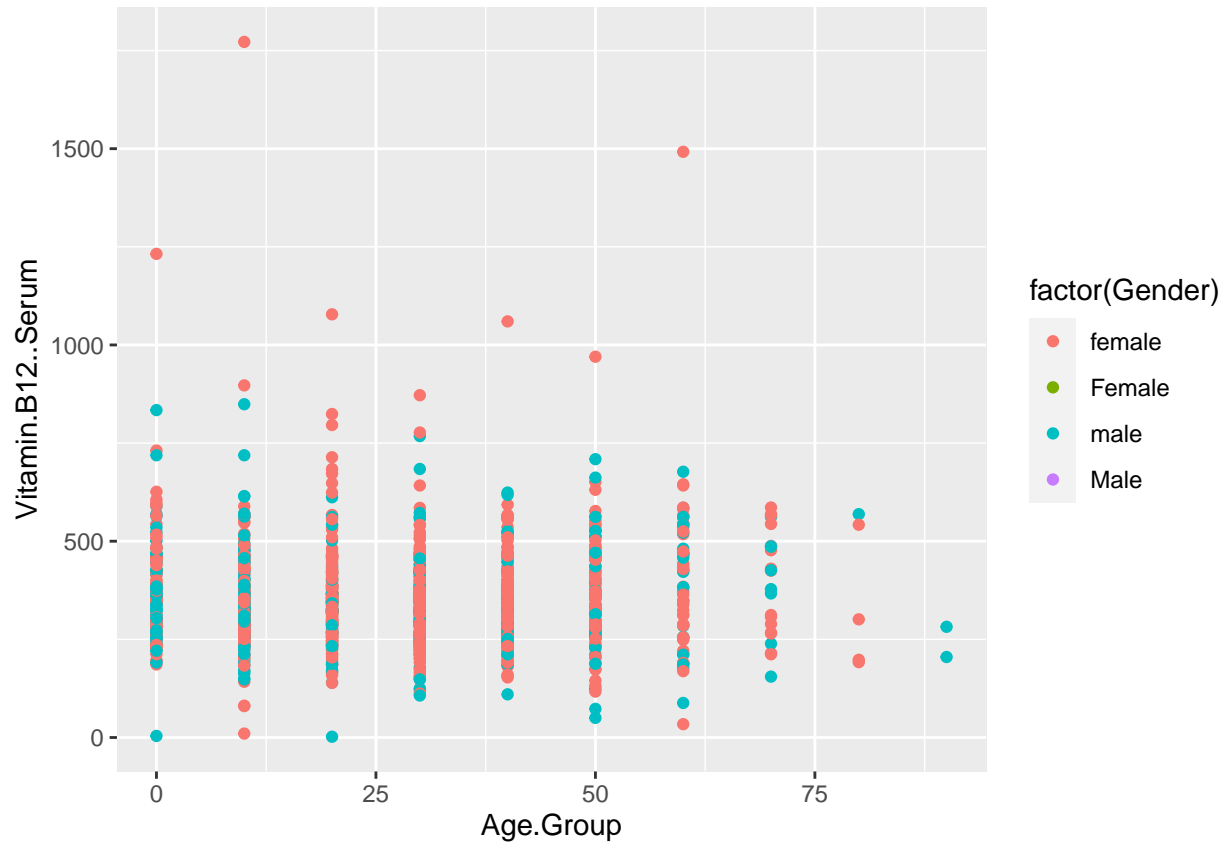
Albumina B12

Data frame with rows with Albumin and Vitamin.B12 values different to NA

```
b12_df <- df_mod %>% filter(!is.na(Albumin...Serum) & !is.na(Vitamin.B12..Serum))
```

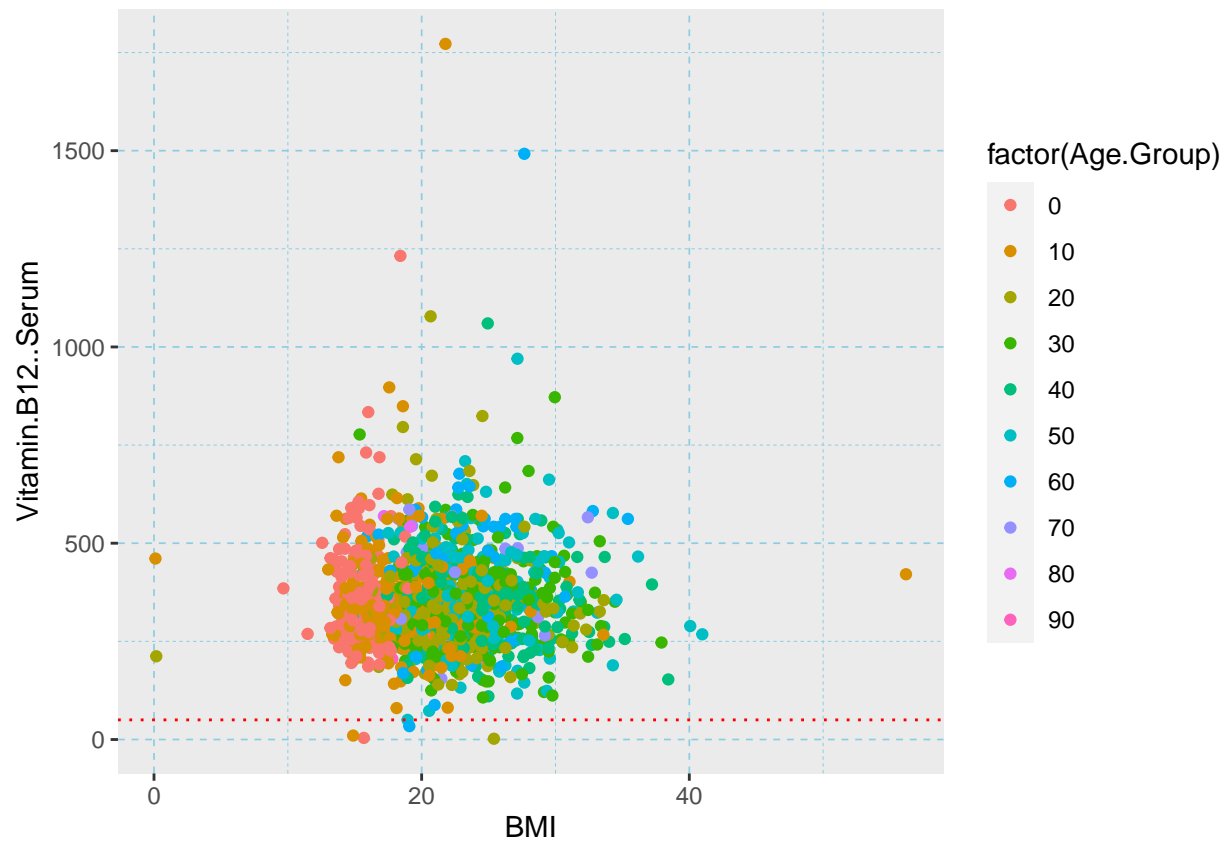
Vitamin.B12 - Age Group Plot

```
b12_df %>% ggplot(aes(x=Age.Group, y=Vitamin.B12..Serum)) + geom_point(aes(colour=factor(Gender)))
```



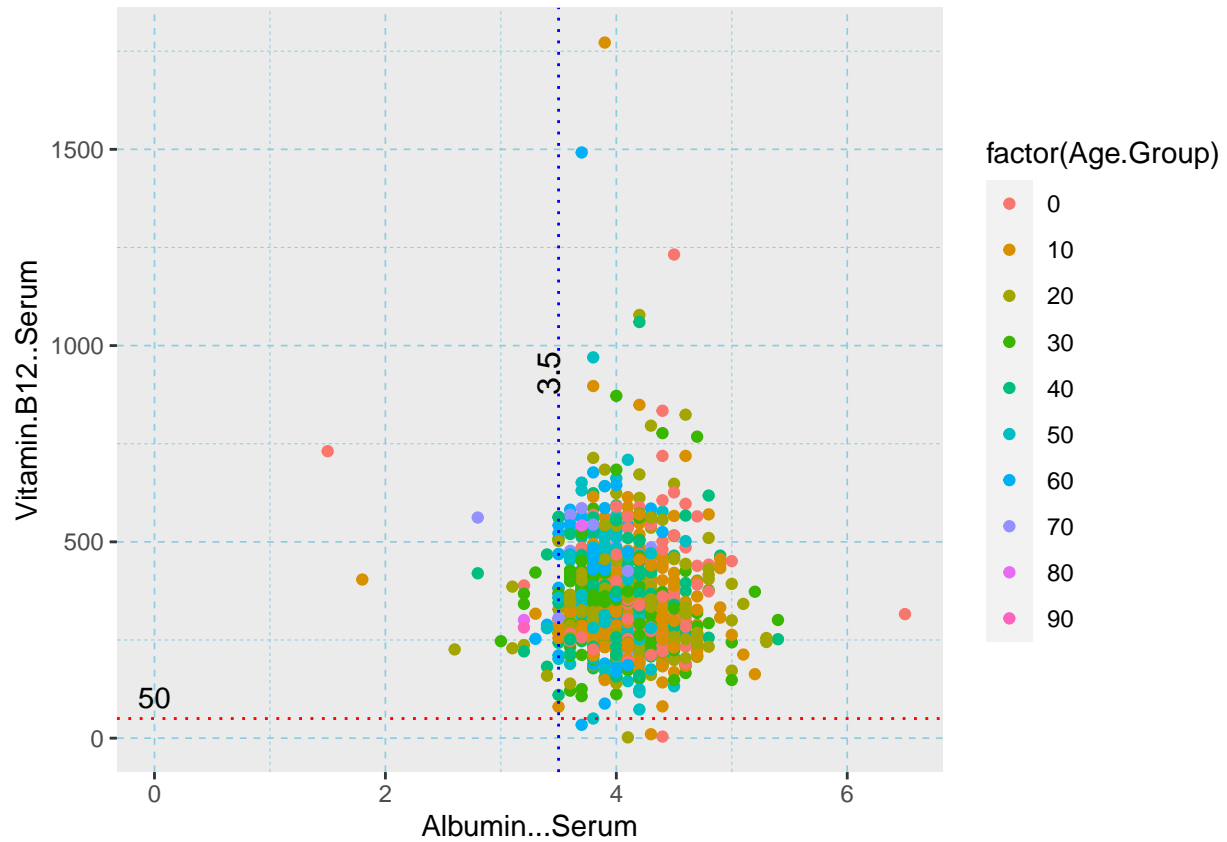
Vitamin B12 - BMI Plot

```
b12_df %>% ggplot(aes(x=BMI, y=Vitamin.B12..Serum)) + geom_point(aes(colour=factor(Age.Group))) + theme_minimal()
```



Albumin - Vitamin B12 group by Age scatter plot

```
b12_df %>% ggplot(aes(x=Albumin...Serum, y = Vitamin.B12..Serum)) + geom_point(aes(colour=factor(Age.Group))) +  
  theme(panel.grid = element_line(color = '#8ccde3', size = 0.3, linetype = 2)) +  
  geom_hline(yintercept = 50, linetype = 'dotted', col = 'red') + annotate("text", x = 0, y = 50, label = "50") +  
  geom_vline(xintercept = 3.5, linetype = 'dotted', col = 'blue') + annotation_custom(grid::textGrob("3.5",
```



Detail data of people with Vitamin.B12 < 50

```
b12_df %>% filter(Vitamin.B12..Serum < 50.0 ) %>% select(Albumin...Serum, Vit.D.assay, Gender, Village
```

##	Albumin...Serum	Vit.D.assay	Gender	Village	Age.Group
## 1	4.4	20	male	Lumdaitkhla	0
## 2	4.3	22	female	Umden Umsaitprah	10
## 3	3.7	39	female	Palwi	60
## 4	4.1	21	male	Umshaproh	20

##	Family.Id
## 1	SVM00000000000130
## 2	SVM00000000000222
## 3	SVM00000000000319
## 4	SVM00000000000343

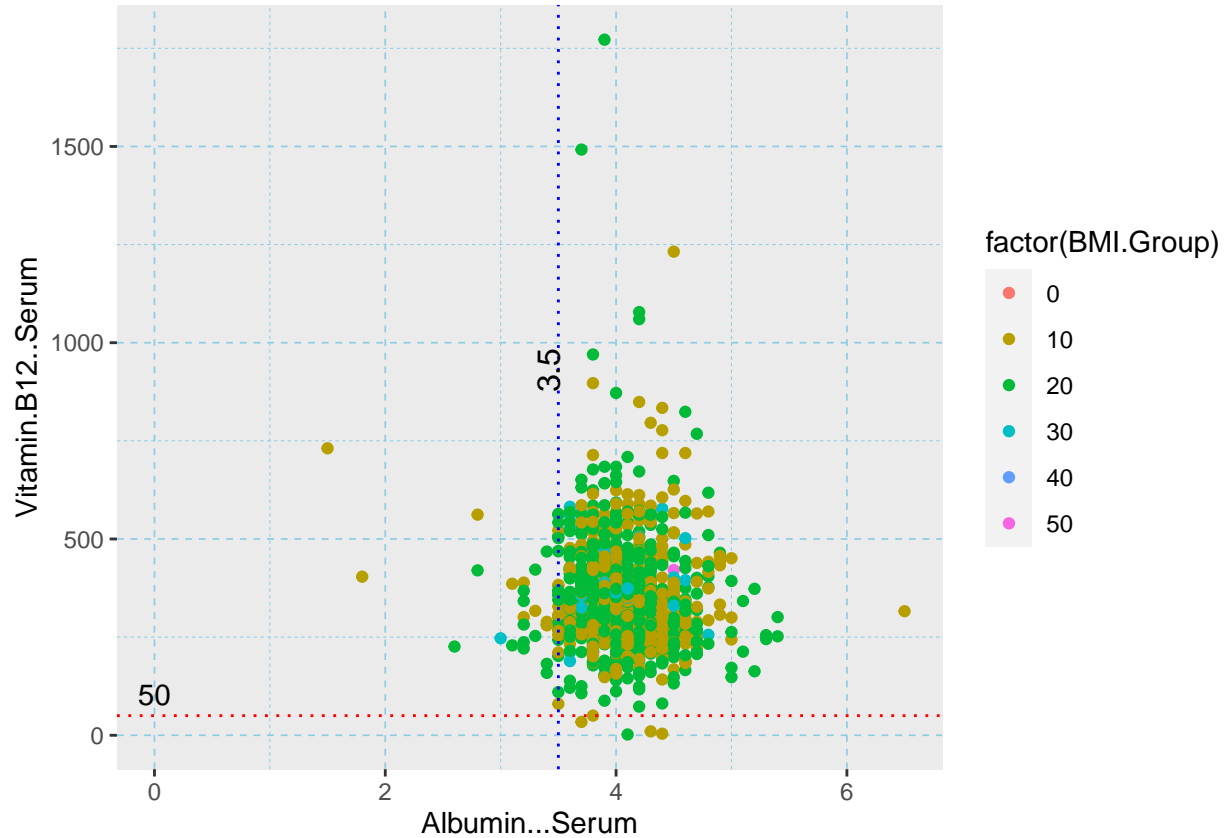
Detail data of people with Vitamin B12 < 50 and Albumin < 3.5

```
b12_df %>% filter(Vitamin.B12..Serum< 5.0 & Albumin...Serum < 3.5 ) %>% select(Albumin...Serum, Vit.D..
```

```
## [1] Albumin...Serum    Vit.D.assay          Vitamin.B12..Serum Gender
## [5] Village              Age.Group            Family.Id
## <0 rows> (or 0-length row.names)
```

Albumin - Vitamin.B12 and BMI group plot

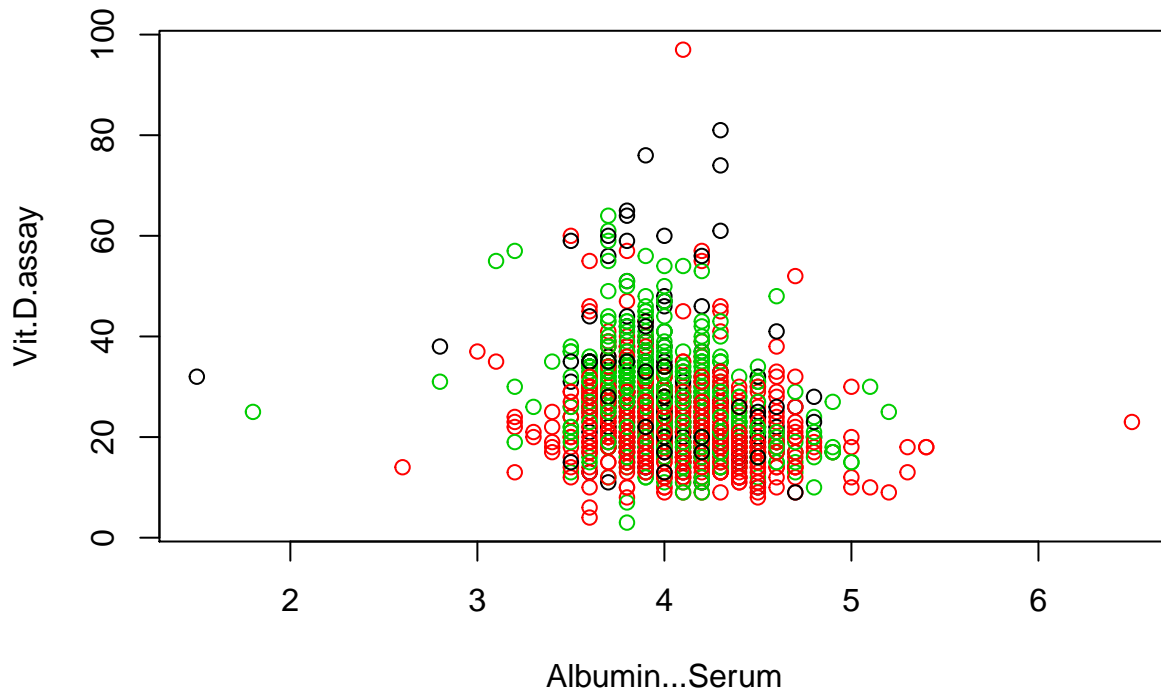
```
b12_df %>% ggplot(aes(x=Albumin...Serum, y = Vitamin.B12..Serum)) + geom_point(aes(colour=factor(BMI.Group))) +  
  geom_hline(yintercept = 50, linetype = 'dotted', col = 'red') + annotate("text", x = 0, y = 50, label = "50") +  
  geom_vline(xintercept = 3.5, linetype = 'dotted', col = 'blue') + annotation_custom(grid::textGrob("3.5",
```




```

## [1074] 3 3 3 2 3 2 2 2 2 3 2 2 2 2 2 3 2 2 2 3 3 3 3 2 2 3 3 3 1 3 2 3 2 2 1 2
## [1111] 3 3 3 2 1 1 1 1 3 3 2 3 3 3 3 3 1 3 3 2 2 2 2 3 2 1 2 3 3 2 3 2 1 1 3 2 2
## [1148] 2 2 2 2 3 3 1 2 3 2 2 3 2 2 2 2 3 2 2 2 2 2 2 1 2 1 2 2 1 2 2 3 1 2 1 1
## [1185] 2 2 2 2 2 2 2 2 2 3 3 2 2 2 2 2 3 3 3 2 2 3 2 2 2 2 2 2 3 2 2 2 2 3 2 3 3
## [1222] 2 2 2 3 2 1 2 3 1 1 3 2 2 3 2 3 2 2 3 2 1 3 3 3 2 2 3 3 3 1 2 2 3 2 2 3 1
## [1259] 2 1 3 2 2 3 3 3 3 2 2 2 2 3 3 2 2 3 3 2 2 2 3 2 2 2 2 2 2 2 2 2 3 1 2 3 3
## [1296] 3 3 3 2 2 2 1 2 3 2 2 2 2 2 3 2 2 2 2 1 2 3 2 2 2 3 2 2 2 2 3 3 2 2 2 2 1
## [1333] 3 2 2 2 3 2 3 3 2 3 3 3 2 2 3 2 2 2 2 2 2 3 1 3 3 3 2 3 2 1 3 3 1 2 2 2 3
## [1370] 3 2 3 1 2 2 2 2 2 3 3 2 2 2 2 1 3 2 1 1 1 1 2 3 2 2 2 2 2 3 3 2 3 3 1 2 1
##
## Within cluster sum of squares by cluster:
## [1] 3974837 1908696 1329456
## (between_SS / total_SS = 69.4 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"

```



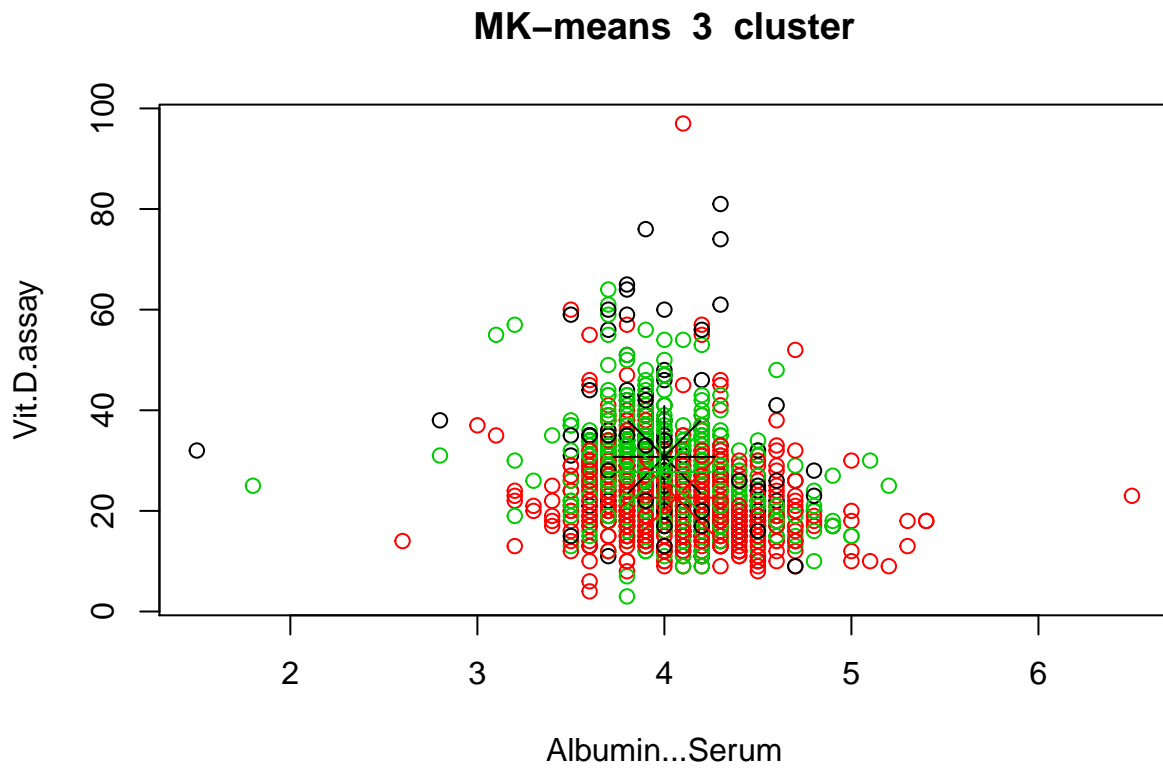
```

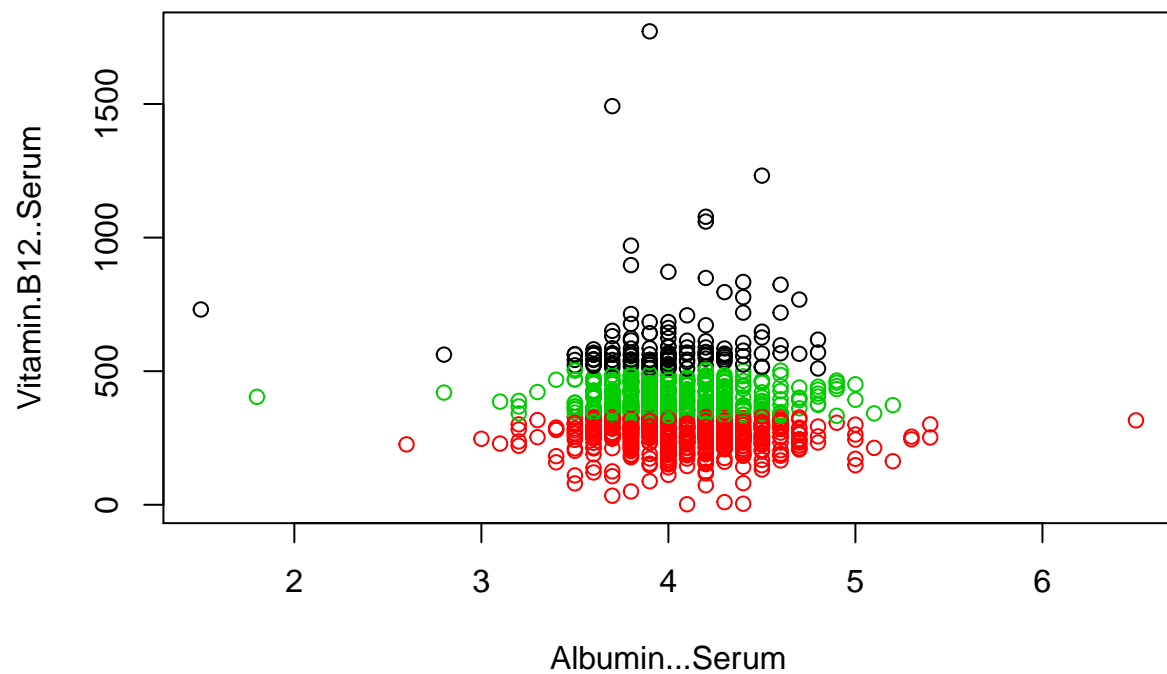
## Albumin...Serum Vit.D.assay Vitamin.B12..Serum
## 1 3.999259 30.73333 616.8889
## 2 4.063556 22.62370 261.2430
## 3 3.976678 27.72987 399.3557

## Albumin...Serum Vit.D.assay
## 1 3.999259 30.73333
## 2 4.063556 22.62370

```

3 3.976678 27.72987

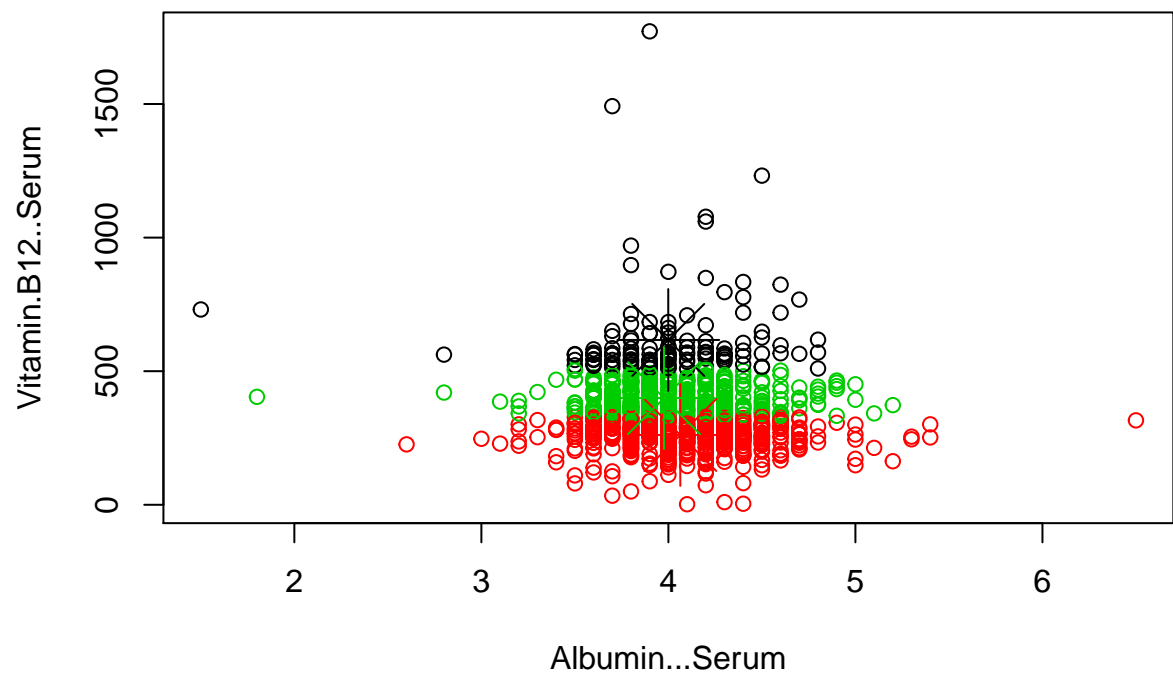




```
## Albumin...Serum Vit.D.assay Vitamin.B12..Serum
## 1      3.999259      30.73333      616.8889
## 2      4.063556      22.62370      261.2430
## 3      3.976678      27.72987      399.3557

## Albumin...Serum Vitamin.B12..Serum
## 1      3.999259      616.8889
## 2      4.063556      261.2430
## 3      3.976678      399.3557
```

MK-means 3 cluster



3D scatter plot

