R assignment

NAMARA IRENE

6/23/2021

call packages required for analysis

1.Import data in R and provide statistics

```
#read csv file diabimmune_16s_t1d_metadata.csv into variable diabetes_data
diabetes_data <- read_csv("diabimmune_16s_t1d_metadata.csv")
```

```
## Parsed with column specification:
## cols(
##   Sample_ID = col_character(),
##   Subject_ID = col_character(),
##   Case_Control = col_character(),
##   Gender = col_character(),
##   Delivery_Route = col_character(),
##   Age_at_Collection = col_integer()
## )
```

```
#view data
view(diabetes_data)
# Change structure of categorical variables to factors
diabetes_data$Gender <- as.factor(diabetes_data$Gender)
diabetes_data$Case_Control <- as.factor(diabetes_data$Case_Control)
diabetes_data$Delivery_Route <- as.factor(diabetes_data$Delivery_Route)

# statistical summary
summary(diabetes_data)
```

```
##    Sample_ID          Subject_ID         Case_Control    Gender
##  Length:777         Length:777         case   :260    female:412
##  Class :character   Class :character   control:517    male  :365
##  Mode  :character   Mode  :character
##
##
##
##   Delivery_Route Age_at_Collection
##  cesarian: 66    Min.   :   6.0
##  vaginal :711    1st Qu.: 229.0
##                  Median : 452.0
##                  Mean   : 482.9
##                  3rd Qu.: 702.0
##                  Max.   :1233.0
```

```
#summarise by counts/frequencies gender and casecontrol
my_count <- xtabs(~ Case_Control + Gender, data = diabetes_data)
my_count
```

```
##             Gender
## Case_Control female male
##       case      142  118
##       control   270  247
```

```
# 142 females and 118 males as cases while 270 females and 247 males were controls.
#find proportions in the data.
total.proportions <- prop.table(my_count)
total.proportions
```

```
##             Gender
## Case_Control    female      male
##       case    0.1827542 0.1518662
##       control 0.3474903 0.3178893
```

```
row.proportions <- prop.table(my_count,1)
row.proportions
```

```
##             Gender
## Case_Control    female      male
##       case    0.5461538 0.4538462
##       control 0.5222437 0.4777563
```

```
column.proportions <- prop.table(my_count,2)
column.proportions
```

```
##             Gender
## Case_Control    female      male
##       case    0.3446602 0.3232877
##       control 0.6553398 0.6767123
```

```
#summary with percentages
summarised_table <- table1(~ Gender + Delivery_Route + Age_at_Collection | Case_Control, data=diabetes_
summarised_table
```

```
## [1] "<table class=\"Rtable1\">\n<thead>\n<tr>\n<th class='rowlabel firstrow lastrow'></th>\n<th class
```

Description

Data as described using the variables provided in the data, Proportions were calculated and percentages obtained as shown in the summarised table.53% of the participants in the study were females and 47% were males.142 females and 118 males were cases while 270 females and 247 males were controls. All cases in the study were born through the vaginal delivery route, 87.2% of the controls in the study were born through the vaginal delivery route. Mean age at collection of the cases was 498 days with a standard deviation of 302. These are further seen in the plots shown below.
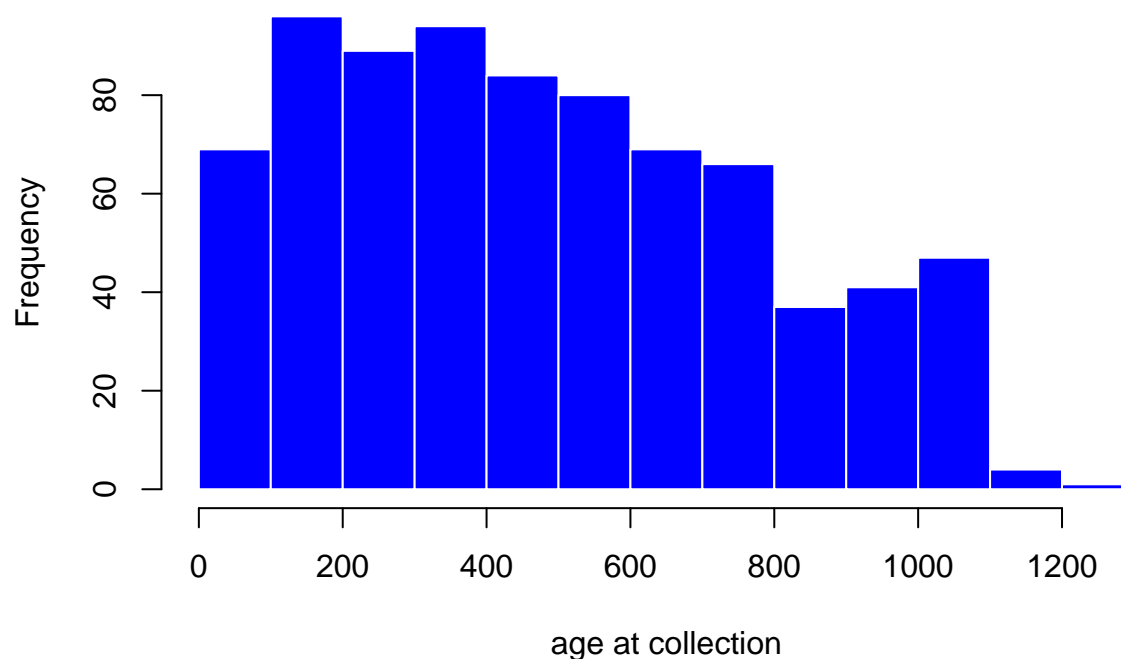
Including Plots for graphics and data visualisation

```
#plot 1:simple histogram
histogram1 <- hist(diabetes_data$Age_at_Collection,xlab = "age at collection",main = "Histogram of Age
```

Histogram of Age at collection
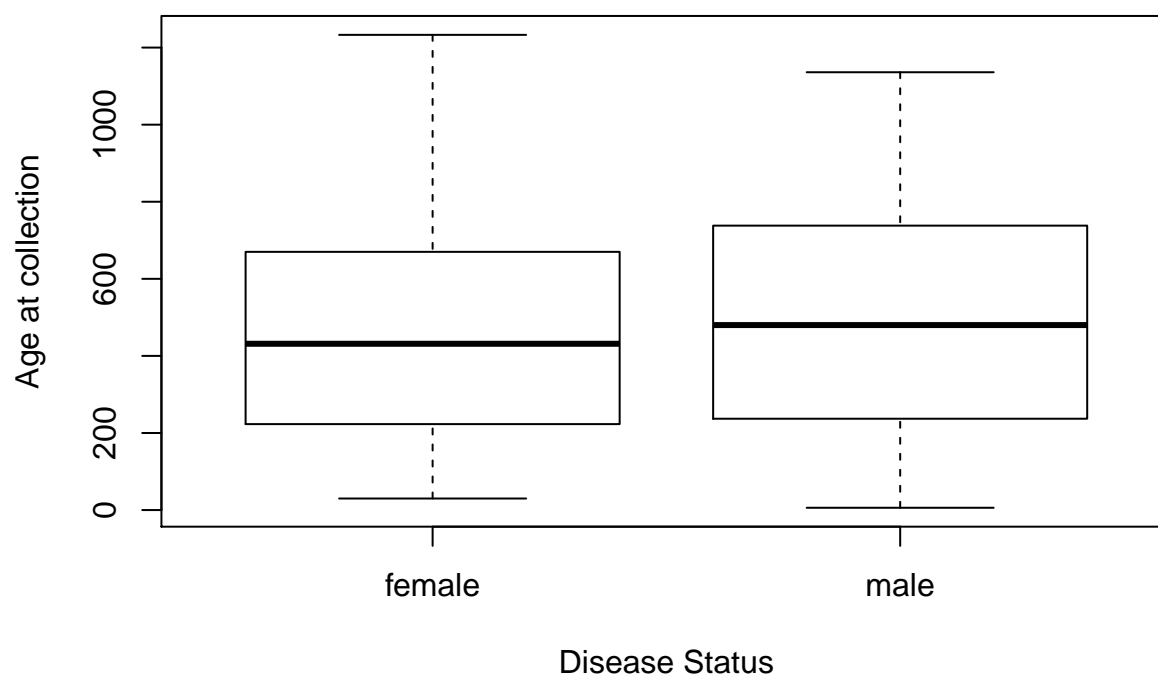


```
histogram1
```

```
## $breaks
##  [1]    0  100  200  300  400  500  600  700  800  900 1000 1100 1200 1300
##
## $counts
##  [1] 69 96 89 94 84 80 69 66 37 41 47  4  1
##
## $density
##  [1] 8.880309e-04 1.235521e-03 1.145431e-03 1.209781e-03 1.081081e-03
##  [6] 1.029601e-03 8.880309e-04 8.494208e-04 4.761905e-04 5.276705e-04
## [11] 6.048906e-04 5.148005e-05 1.287001e-05
##
## $mids
##  [1]   50  150  250  350  450  550  650  750  850  950 1050 1150 1250
##
## $xname
## [1] "diabetes_data$Age_at_Collection"
##
## $equidist
## [1] TRUE
##
## attr(,"class")
## [1] "histogram"
```

```r
#plot2:simple boxplot
boxplot1 <- boxplot(diabetes_data$Age_at_Collection~ diabetes_data$Gender, main="A simple boxplot showi
```

```
        ylab="Age at collection")
```

A simple boxplot showing Age at Collection by Disease Status


```
boxplot1
```

```
## $stats
##        [,1] [,2]
## [1,]   30.0    6
## [2,]  223.0  237
## [3,]  431.5  480
## [4,]  670.0  738
## [5,] 1233.0 1136
## attr(,"class")
##    female
## "integer"
##
## $n
## [1] 412 365
##
## $conf
##          [,1]     [,2]
## [1,] 396.7051 438.5668
## [2,] 466.2949 521.4332
##
## $out
## numeric(0)
##
```

```
## $group
## numeric(0)
##
## $names
## [1] "female" "male"
```
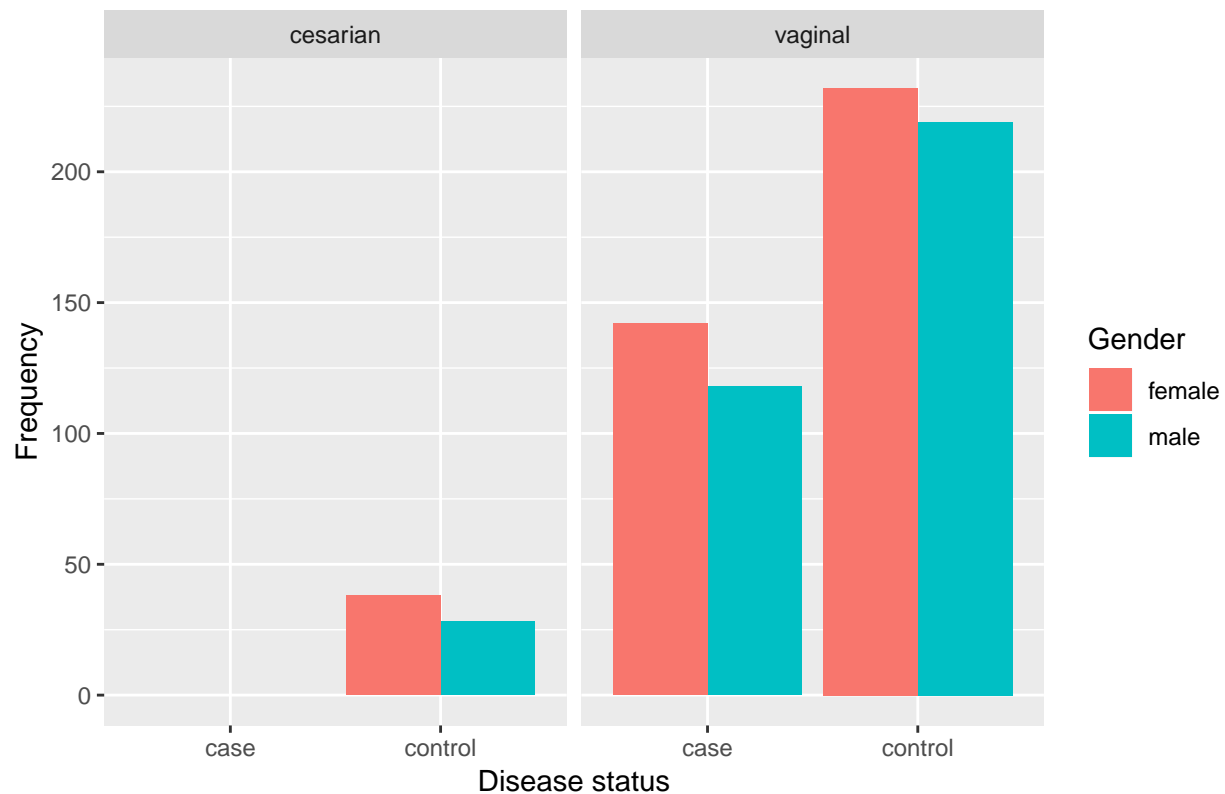
```r
#ggplots
library(ggplot2)
#plot3: ggplot to compare the categorical variables
ggplot_1 <- ggplot(diabetes_data,aes(x=Case_Control))+ geom_bar(aes(fill=Gender), position = "dodge") +
ggplot_1
```

Comparison of the categorical variables


```r
#plot4:ggplot for numerical variable, age at collection
ggplot_2 <- ggplot(data=diabetes_data) + geom_histogram(aes(x=Age_at_Collection),color="red3")+
  scale_x_continuous(name="Age_at_Collection") + ggtitle(label="Histogram of Age at Collection")+
theme_classic() + scale_fill_brewer(palette = "Spectral")
ggplot_2
```

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

Histogram of Age at Collection

Test for associations/independency

```r
# option1:compare disease status with gender and Delivery_Route using chisquare test.
#Gender
tab1 <- table(diabetes_data$Case_Control,diabetes_data$Gender)
tab1
```

```
##
##            female male
##    case       142  118
##    control    270  247
```

```r
chisquaretest1 <- chisq.test(tab1)
chisquaretest1
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  tab1
## X-squared = 0.30687, df = 1, p-value = 0.5796
```

```r
#Delivery_route
chisq.test(tab1)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
```

```
##
## data:  tab1
## X-squared = 0.30687, df = 1, p-value = 0.5796
```

```
tab2 <- table(diabetes_data$Case_Control,diabetes_data$Delivery_Route)
tab2
```

```
##
##            cesarian vaginal
##    case           0     260
##    control       66     451
```

```
chisquaretest2 <- chisq.test(tab2)
chisquaretest2
```

```
##
##   Pearson's Chi-squared test with Yates' continuity correction
##
## data:  tab2
## X-squared = 34.649, df = 1, p-value = 3.949e-09
```

```
#option 2: compare disease status with age at collection using the two sample t-test
my_t.test <- t.test(diabetes_data$Age_at_Collection~diabetes_data$Case_Control)
my_t.test
```

```
##
##   Welch Two Sample t-test
##
## data:  diabetes_data$Age_at_Collection by diabetes_data$Case_Control
## t = 1.0195, df = 502.81, p-value = 0.3085
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -21.43301  67.66969
## sample estimates:
##    mean in group case mean in group control
##              498.2731              475.1547
```

Interpretation

- Disease status is not dependant on the gender of the participant since the p-value obtained after running the first Chi-squared test is greater than 0.01 and therefore insignificant.

- Disease status is dependant on the delivery route of the participant because p-value that was obtained after running the second Chi-squared test is less than 0.01 and is significant.

- Disease status is not associated with the age at collection of the participant as shown by the insignificant p-value obtained with the t-test done.

2.Using phyloseq, create a phyloseq object.

```
#Refine and import data to create Phyloseq Object
phyloseq_diabetes_data <- diabetes_data #assign diabetes_data a new variable phyloseq_diabetes_data
phyloseq_diabetes_data <- data.frame(phyloseq_diabetes_data,row.names = NULL)
Taxonomysummary <- read_excel("Taxonomysummary.xlsx")
```

```r
OTU_TABLE <- read_excel("OTU_TABLE.xlsx")
#view first few lines of the new variables formed.
head(phyloseq_diabetes_data)
```

```
##   Sample_ID Subject_ID Case_Control Gender Delivery_Route Age_at_Collection
## 1    G36449    E001463      control   male        vaginal                62
## 2    G36034    E001463      control   male        vaginal                82
## 3    G36993    E001463      control   male        vaginal               124
## 4    G35523    E001463      control   male        vaginal               153
## 5    G36450    E001463      control   male        vaginal               187
## 6    G36028    E001463      control   male        vaginal               213
```

```r
head(Taxonomysummary)
```

```
## # A tibble: 6 x 8
##     taxa1 Kingdom   Phylum     Class      Order     Family    Genus    Species
##     <dbl> <chr>     <chr>      <chr>      <chr>     <chr>     <chr>    <chr>
## 1 4333897 k__Bacte~ p__Proteo~ c__Gammapr~ o__Enter~ f__Entero~ g__     s__
## 2  190162 k__Bacte~ p__Firmic~ c__Clostri~ o__Clost~ f__Lachno~ g__Blau~ s__
## 3  134726 k__Bacte~ p__Firmic~ c__Bacilli  o__Lacto~ f__Lactob~ g__Lact~ s__
## 4  679245 k__Bacte~ p__Firmic~ c__Bacilli  o__Lacto~ f__Lactob~ g__Lact~ s__
## 5  289734 k__Bacte~ p__Firmic~ c__Clostri~ o__Clost~ f__Lachno~ g__     s__
## 6  302049 k__Bacte~ p__Firmic~ c__Clostri~ o__Clost~ f__Lachno~ g__Blau~ s__
```

```r
head(OTU_TABLE)
```

```
## # A tibble: 6 x 778
##     otu1 G37016 G36918 G37044 G37009 G37029 G37035 G36982 G36984 G37030 G37031
##    <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>
## 1 4.33e6     12      1     12      0      0     22      0      0      0     35
## 2 1.90e5      2      1      0      1      4      0     15     20      3      0
## 3 1.35e5     13    127      0      0      0      0      0      0      0      0
## 4 6.79e5     10      4      0      0      0     15      0      9      0      0
## 5 2.90e5    313    923    101    701    545      0   2786   3756    440     10
## 6 3.02e5     11     40      4     22     20      0    101    190     25      0
## # ... with 767 more variables: G36906 <dbl>, G37028 <dbl>, G37014 <dbl>,
## #   G37010 <dbl>, G36944 <dbl>, G36902 <dbl>, G37051 <dbl>, G36964 <dbl>,
## #   G36951 <dbl>, G37025 <dbl>, G36930 <dbl>, G36935 <dbl>, G37041 <dbl>,
## #   G36959 <dbl>, G36905 <dbl>, G36917 <dbl>, G36992 <dbl>, G36921 <dbl>,
## #   G37004 <dbl>, G36966 <dbl>, G36936 <dbl>, G37015 <dbl>, G37034 <dbl>,
## #   G36953 <dbl>, G36974 <dbl>, G37039 <dbl>, G36933 <dbl>, G37052 <dbl>,
## #   G36925 <dbl>, G37046 <dbl>, G37053 <dbl>, G36998 <dbl>, G37045 <dbl>,
## #   G37049 <dbl>, G37019 <dbl>, G36988 <dbl>, G37037 <dbl>, G37005 <dbl>,
## #   G36928 <dbl>, G36923 <dbl>, G37033 <dbl>, G37017 <dbl>, G37012 <dbl>,
## #   G36952 <dbl>, G37011 <dbl>, G36957 <dbl>, G37021 <dbl>, G36937 <dbl>,
## #   G37047 <dbl>, G37032 <dbl>, G36932 <dbl>, G37043 <dbl>, G37054 <dbl>,
## #   G36948 <dbl>, G36972 <dbl>, G36968 <dbl>, G36911 <dbl>, G36987 <dbl>,
## #   G37042 <dbl>, G36960 <dbl>, G37040 <dbl>, G37027 <dbl>, G37024 <dbl>,
## #   G37023 <dbl>, G37036 <dbl>, G37026 <dbl>, G37050 <dbl>, G36973 <dbl>,
## #   G36991 <dbl>, G36904 <dbl>, G36958 <dbl>, G36993 <dbl>, G37020 <dbl>,
## #   G36910 <dbl>, G36927 <dbl>, G37018 <dbl>, G37048 <dbl>, G36940 <dbl>,
## #   G37013 <dbl>, G37038 <dbl>, G36938 <dbl>, G37022 <dbl>, G35534 <dbl>,
## #   G35535 <dbl>, G35361 <dbl>, G35364 <dbl>, G35433 <dbl>, G35528 <dbl>,
## #   G35381 <dbl>, G35416 <dbl>, G35477 <dbl>, G35390 <dbl>, G35536 <dbl>,
## #   G35514 <dbl>, G35525 <dbl>, G35395 <dbl>, G35480 <dbl>, G35409 <dbl>,
```

```
## #   G35391 <dbl>, G35437 <dbl>, ...
```

```r
#Inorder to have similar row names in OTU and taxonomy table, the following was done
phyloseq_diabetes_data <- column_to_rownames(phyloseq_diabetes_data, var = "Sample_ID")
Taxonomysummary <- column_to_rownames(Taxonomysummary, var = "taxa1")
OTU_TABLE <- column_to_rownames(OTU_TABLE, var = "otu1")

# create matrices from dataframes: OTU_table and Taxonomysummary
class(OTU_TABLE)
```

```
## [1] "data.frame"
```

```r
OTU_table_matrix <- as.matrix(OTU_TABLE)
class(OTU_table_matrix)
```
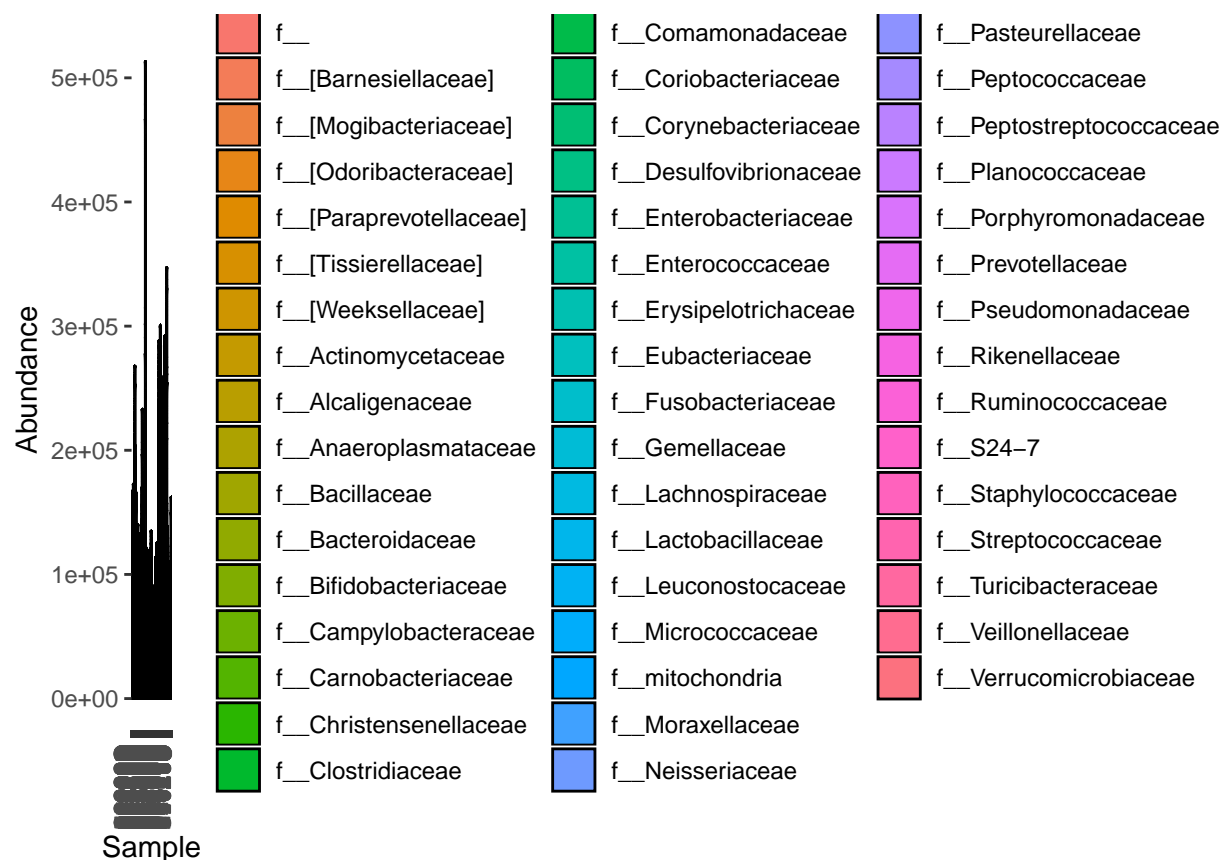
```
## [1] "matrix"
```

```r
taxonomymatrix <- as.matrix(Taxonomysummary)
class(taxonomymatrix)
```

```
## [1] "matrix"
```

```r
# create phyloseq object
my_OTUtable <-  otu_table(OTU_table_matrix, taxa_are_rows = TRUE)
my_Taxonomytable <-  tax_table(taxonomymatrix)
my_sampledata1 <-  sample_data(phyloseq_diabetes_data)
my_phyloseq_object = phyloseq(my_OTUtable, my_Taxonomytable, my_sampledata1)

my_phyloseq_object
```

```
## phyloseq-class experiment-level object
## otu_table()   OTU Table:         [ 2240 taxa and 777 samples ]
## sample_data() Sample Data:       [ 777 samples by 5 sample variables ]
## tax_table()   Taxonomy Table:    [ 2240 taxa by 7 taxonomic ranks ]
```

```r
#visualization and checking validity of phyloseq object created
plot_bar(my_phyloseq_object, fill = "Family")
```

Abundance (y-axis): 5e+05, 4e+05, 3e+05, 2e+05, 1e+05, 0e+00

Sample (x-axis)

Legend:

f__
f__[Barnesiellaceae]
f__[Mogibacteriaceae]
f__[Odoribacteraceae]
f__[Paraprevotellaceae]
f__[Tissierellaceae]
f__[Weeksellaceae]
f__Actinomycetaceae
f__Alcaligenaceae
f__Anaeroplasmataceae
f__Bacillaceae
f__Bacteroidaceae
f__Bifidobacteriaceae
f__Campylobacteraceae
f__Carnobacteriaceae
f__Christensenellaceae
f__Clostridiaceae
f__Comamonadaceae
f__Coriobacteriaceae
f__Corynebacteriaceae
f__Desulfovibrionaceae
f__Enterobacteriaceae
f__Enterococcaceae
f__Erysipelotrichaceae
f__Eubacteriaceae
f__Fusobacteriaceae
f__Gemellaceae
f__Lachnospiraceae
f__Lactobacillaceae
f__Leuconostocaceae
f__Micrococcaceae
f__mitochondria
f__Moraxellaceae
f__Neisseriaceae
f__Pasteurellaceae
f__Peptococcaceae
f__Peptostreptococcaceae
f__Planococcaceae
f__Porphyromonadaceae
f__Prevotellaceae
f__Pseudomonadaceae
f__Rikenellaceae
f__Ruminococcaceae
f__S24–7
f__Staphylococcaceae
f__Streptococcaceae
f__Turicibacteraceae
f__Veillonellaceae
f__Verrucomicrobiaceae

```
rank_names(my_phyloseq_object)
```

```
## [1] "Kingdom" "Phylum" "Class"   "Order"   "Family"  "Genus"   "Species"
```

```
taxa_names(my_phyloseq_object)
```

```
##     [1] "4333897" "190162"  "134726"  "679245"  "289734"  "302049"  "197991"
##     [8] "3903651" "184922"  "193946"  "189592"  "3211875" "175261"  "197970"
##    [15] "4390365" "186624"  "4366867" "175612"  "175184"  "4454531" "198422"
##    [22] "4448492" "194101"  "730906"  "1740283" "172777"  "3583645" "355975"
##    [29] "189920"  "798581"  "193672"  "3275562" "4391262" "2762219" "176067"
##    [36] "184990"  "4439360" "1504042" "306299"  "196957"  "2017729" "3141094"
##    [43] "4446902" "1142029" "4354235" "198646"  "941096"  "1076969" "536866"
##    [50] "197930"  "782953"  "174818"  "176306"  "3472078" "1992"    "4372612"
##    [57] "593043"  "2565100" "4425571" "179826"  "184037"  "365621"  "4481861"
##    [64] "197072"  "192684"  "368490"  "2700883" "198909"  "175729"  "308544"
##    [71] "4424598" "4375664" "271159"  "4385326" "344525"  "192958"  "3926480"
##    [78] "365536"  "194177"  "180999"  "4316928" "4480189" "186687"  "196990"
##    [85] "72820"   "342397"  "197552"  "4297695" "197364"  "577294"  "2544615"
##    [92] "173863"  "825808"  "4401580" "4060124" "198183"  "4372625" "187447"
##    [99] "187569"  "197481"  "4342104" "196246"  "4395096" "2683271" "4437515"
##   [106] "535955"  "174672"  "1860112" "4364745" "198947"  "3943182" "4419459"
##   [113] "292735"  "313593"  "3134492" "4020046" "182289"  "59563"   "182886"
##   [120] "656881"  "4381553" "197022"  "1141218" "1508541" "340876"  "676211"
##   [127] "349431"  "225919"  "212532"  "4278525" "190490"  "403701"  "4467447"
##   [134] "3801267" "12574"   "2920309" "3154070" "199702"  "4433947" "198071"
##   [141] "2157225" "192127"  "577170"  "302746"  "1839271" "523140"  "2876801"
```

```
## [148] "588471"  "180509"  "175485"  "4469576" "4435664" "198209"  "189450"
## [155] "4103317" "3272764" "198930"  "312882"  "3537197" "4426298" "187126"
## [162] "179961"  "223059"  "2283862" "191361"  "4416025" "191081"  "174288"
## [169] "199694"  "162491"  "4372382" "184209"  "192308"  "4439469" "4339832"
## [176] "174862"  "851715"  "188079"  "3327894" "192536"  "4438116" "193528"
## [183] "192406"  "4449054" "180462"  "4372528" "174763"  "526804"  "195651"
## [190] "180826"  "189997"  "146554"  "4458959" "170462"  "4329132" "3907189"
## [197] "177111"  "4468234" "181139"  "4325509" "184760"  "587933"  "291090"
## [204] "211706"  "178238"  "4315974" "4473788" "4425214" "367176"  "187882"
## [211] "187883"  "194151"  "4414476" "4472399" "340474"  "3805518" "199145"
## [218] "174010"  "178439"  "4144205" "188735"  "199761"  "4457427" "193233"
## [225] "3439402" "187945"  "3931537" "2714942" "4436046" "190502"  "4453501"
## [232] "192444"  "180037"  "180216"  "2497335" "4094866" "175535"  "198449"
## [239] "4256470" "493391"  "211935"  "4450010" "4442130" "328472"  "3887769"
## [246] "161423"  "4393532" "4483015" "186888"  "4424239" "192383"  "3600504"
## [253] "4309301" "305318"  "199293"  "195508"  "185607"  "197004"  "471180"
## [260] "195465"  "180606"  "175751"  "4477696" "190058"  "716984"  "190301"
## [267] "4381338" "178799"  "183604"  "199430"  "4183249" "177581"  "182073"
## [274] "191601"  "4464173" "4420408" "4411295" "3393191" "322114"  "194686"
## [281] "4447072" "187077"  "2415144" "180414"  "196664"  "3709990" "259772"
## [288] "578016"  "54794"   "553611"  "48084"   "1751298" "179657"  "4396426"
## [295] "4472091" "324894"  "4378683" "197760"  "4383953" "4434334" "2575651"
## [302] "4451899" "2549971" "4439603" "328617"  "365385"  "2986828" "4437814"
## [309] "358798"  "4347159" "179200"  "189208"  "359314"  "3251419" "2979308"
## [316] "181864"  "4331782" "2724175" "4481613" "176753"  "356745"  "4306048"
## [323] "190761"  "4306262" "3475269" "182383"  "197490"  "184925"  "180563"
## [330] "288810"  "326482"  "4465907" "1830364" "357261"  "4357811" "4457268"
## [337] "199054"  "2309802" "3265161" "162623"  "180123"  "295146"  "180552"
## [344] "176967"  "187210"  "335577"  "4419650" "348304"  "179911"  "193075"
## [351] "367456"  "194899"  "4473596" "4344860" "3039313" "363735"  "193887"
## [358] "4232045" "4390319" "190100"  "4111715" "4379957" "186921"  "4335781"
## [365] "4381555" "177175"  "193667"  "197334"  "187952"  "2061739" "193174"
## [372] "2298935" "332185"  "176604"  "292521"  "189630"  "183970"  "194659"
## [379] "2148365" "4439690" "4404220" "4004991" "177561"  "191874"  "191872"
## [386] "175704"  "182911"  "4413347" "185411"  "1848900" "180352"  "3443092"
## [393] "173135"  "2704013" "177237"  "174614"  "4383922" "194053"  "207390"
## [400] "4445673" "2177184" "182506"  "179472"  "3531225" "183496"  "180289"
## [407] "307113"  "231952"  "198740"  "350832"  "186263"  "4370024" "2654263"
## [414] "191595"  "197864"  "189647"  "177754"  "3195500" "180528"  "180529"
## [421] "3450454" "208739"  "188127"  "362765"  "191399"  "1096610" "354850"
## [428] "206574"  "3413566" "177222"  "593422"  "194670"  "1917420" "197458"
## [435] "851938"  "304047"  "4079463" "182245"  "302932"  "196307"  "2072645"
## [442] "4354477" "340113"  "3943186" "192711"  "216111"  "193666"  "3576174"
## [449] "185235"  "350865"  "1145098" "4456252" "4345850" "170652"  "191928"
## [456] "182054"  "199677"  "178763"  "141145"  "193477"  "189271"  "183662"
## [463] "177986"  "787663"  "183030"  "184864"  "16513"   "567960"  "3014082"
## [470] "4383924" "189150"  "178015"  "2250985" "4375000" "193968"  "193763"
## [477] "4447950" "199190"  "184910"  "4340358" "192735"  "186955"  "1654474"
## [484] "176057"  "681370"  "349257"  "267718"  "185192"  "4326080" "1142110"
## [491] "199279"  "183849"  "1085832" "195929"  "178151"  "4373152" "198956"
## [498] "190980"  "187782"  "182517"  "363029"  "4452633" "196518"  "4422456"
## [505] "192461"  "4094259" "194758"  "3713805" "2689396" "191442"  "3203801"
## [512] "2656868" "177515"  "4337755" "566243"  "179905"  "287514"  "183340"
## [519] "195015"  "2132002" "199490"  "842594"  "4444262" "311820"  "360660"
```

11

```
## [526] "182643"  "198198"  "1100972" "4377328" "4045882" "358008"  "181754"
## [533] "3244992" "173996"  "299837"  "4388775" "334336"  "189403"  "181899"
## [540] "3588390" "182874"  "318970"  "175550"  "724120"  "566976"  "178462"
## [547] "152485"  "3756485" "355197"  "184477"  "347529"  "3563235" "176865"
## [554] "4426877" "181140"  "3929758" "4296424" "175168"  "196100"  "511378"
## [561] "317677"  "3633129" "1146349" "955102"  "4432433" "864465"  "176300"
## [568] "299267"  "185420"  "1928156" "179267"  "176244"  "2123717" "189782"
## [575] "175145"  "184897"  "28914"   "193367"  "3265556" "2686384" "191214"
## [582] "210542"  "183698"  "189708"  "340547"  "187051"  "964220"  "3715618"
## [589] "579851"  "181560"  "2840201" "210371"  "3857426" "174497"  "3799784"
## [596] "301578"  "4326219" "178885"  "538000"  "4455308" "4323124" "195646"
## [603] "255367"  "270382"  "352347"  "1046997" "4376828" "191043"  "306704"
## [610] "178928"  "187623"  "4343631" "138389"  "334459"  "181155"  "342110"
## [617] "192625"  "4451251" "157327"  "337538"  "175560"  "360636"  "180468"
## [624] "188930"  "177349"  "2318497" "1062051" "661266"  "3439403" "265871"
## [631] "245625"  "392887"  "357046"  "174840"  "2365946" "2617854" "176704"
## [638] "45363"   "3028273" "188008"  "328105"  "4473295" "192046"  "4332082"
## [645] "3794053" "316732"  "188856"  "186478"  "358939"  "3424669" "178001"
## [652] "290281"  "212698"  "193551"  "189205"  "196219"  "4307484" "163243"
## [659] "187385"  "209760"  "4301298" "196315"  "194383"  "192963"  "3225199"
## [666] "4329571" "194104"  "146935"  "3090117" "1919007" "193644"  "3321163"
## [673] "181422"  "4331360" "187212"  "192720"  "2235671" "181756"  "365717"
## [680] "186997"  "186723"  "177520"  "4339831" "190076"  "668514"  "180171"
## [687] "181485"  "4353757" "84709"   "4433823" "3430935" "3244896" "4430843"
## [694] "187178"  "3924627" "183579"  "359872"  "182052"  "4374302" "300235"
## [701] "182577"  "186732"  "4356065" "354574"  "211907"  "177567"  "180312"
## [708] "321453"  "195436"  "4449236" "539819"  "175844"  "196371"  "3609545"
## [715] "4335815" "175586"  "194660"  "354334"  "4381430" "4474255" "185021"
## [722] "184000"  "186255"  "4343627" "192983"  "4452051" "191395"  "194733"
## [729] "182255"  "182854"  "1950496" "193709"  "194287"  "191483"  "190320"
## [736] "4318671" "2237211" "186866"  "329313"  "4380191" "324163"  "191110"
## [743] "312586"  "179384"  "206233"  "2018038" "366794"  "179677"  "328544"
## [750] "4440670" "1108007" "194534"  "4393568" "195123"  "186077"  "319687"
## [757] "4370025" "4386317" "4346677" "4321810" "3138798" "3544699" "1103978"
## [764] "183207"  "365842"  "4407515" "4448211" "190240"  "174950"  "185563"
## [771] "188648"  "185232"  "185731"  "369014"  "105287"  "176201"  "176297"
## [778] "171559"  "517754"  "1820513" "816702"  "1756274" "174439"  "177040"
## [785] "193534"  "188863"  "192231"  "1918929" "4365130" "875735"  "4326452"
## [792] "176507"  "195716"  "187035"  "193744"  "195681"  "193452"  "3426658"
## [799] "347639"  "186416"  "663649"  "4471251" "198119"  "4430639" "194672"
## [806] "188753"  "4466616" "195805"  "701221"  "2119418" "73609"   "4396235"
## [813] "4479397" "4469359" "339087"  "4436552" "191251"  "184339"  "149614"
## [820] "181918"  "315846"  "197675"  "189459"  "586141"  "525215"  "195004"
## [827] "3450453" "184174"  "2281172" "191412"  "304779"  "189755"  "759349"
## [834] "192424"  "175559"  "1065569" "194824"  "177037"  "526682"  "322580"
## [841] "176269"  "4329112" "340219"  "355424"  "180629"  "186030"  "3851391"
## [848] "208479"  "4339144" "163494"  "302352"  "266210"  "3272632" "183829"
## [855] "849535"  "4396292" "4463108" "4318208" "174516"  "311950"  "176062"
## [862] "545299"  "4476780" "2201995" "4458576" "562244"  "189083"  "191153"
## [869] "4020502" "178211"  "177172"  "3236435" "4372003" "301253"  "4434579"
## [876] "4331760" "580629"  "213566"  "179400"  "753638"  "288651"  "3754778"
## [883] "197988"  "4453609" "4470870" "329693"  "2740953" "192263"  "181539"
## [890] "179583"  "16054"   "1943669" "757622"  "363264"  "132892"  "4431545"
## [897] "4410166" "130103"  "4362300" "195937"  "305141"  "181432"  "208539"
```

```
##  [904] "183932"  "367889"  "4281639" "41229"   "4368216" "2943548" "4412022"
##  [911] "4397092" "215269"  "193873"  "187946"  "198786"  "174831"  "184342"
##  [918] "4395075" "179508"  "177230"  "4379449" "332929"  "174611"  "184525"
##  [925] "181003"  "4473509" "4408801" "190453"  "851704"  "184464"  "2037235"
##  [932] "230578"  "292057"  "2210028" "215670"  "187751"  "174625"  "148279"
##  [939] "195105"  "3082155" "4396297" "4433737" "134265"  "1033413" "181204"
##  [946] "190577"  "186022"  "2424737" "185763"  "4364747" "177100"  "178420"
##  [953] "2831841" "4465072" "187356"  "558839"  "563572"  "975306"  "194654"
##  [960] "196382"  "178485"  "167373"  "192015"  "194129"  "296442"  "357449"
##  [967] "1105984" "308333"  "4217963" "196296"  "306499"  "180083"  "346938"
##  [974] "180155"  "184465"  "4462107" "772282"  "184678"  "4420669" "4344207"
##  [981] "2280817" "135956"  "270094"  "885437"  "4413447" "2058521" "198128"
##  [988] "189610"  "198696"  "186090"  "1000547" "619140"  "4483337" "4421070"
##  [995] "4295707" "360000"  "4396688" "188676"  "194586"  "360730"  "4316391"
## [1002] "189860"  "189092"  "193053"  "307238"  "190676"  "292921"  "194667"
## [1009] "195186"  "4416763" "199448"  "176113"  "362568"  "1868703" "264967"
## [1016] "2160415" "4472130" "4331364" "193863"  "1602805" "3747551" "312969"
## [1023] "196054"  "4401375" "4410401" "4417708" "4318122" "186468"  "213394"
## [1030] "174752"  "366147"  "111135"  "185859"  "331253"  "196139"  "2170530"
## [1037] "571642"  "561171"  "342380"  "187780"  "166637"  "184238"  "2532173"
## [1044] "4449055" "182167"  "179695"  "4437024" "553080"  "589329"  "1835779"
## [1051] "192419"  "4443574" "258550"  "3562626" "1061772" "194089"  "191792"
## [1058] "362608"  "4458306" "2438203" "196200"  "198587"  "186549"  "185113"
## [1065] "4432431" "188832"  "923098"  "1736067" "182089"  "4391625" "192676"
## [1072] "4318125" "4418787" "182196"  "4309636" "355510"  "3318103" "147969"
## [1079] "180874"  "4323555" "4301511" "195556"  "1607319" "178760"  "191180"
## [1086] "4347520" "866280"  "279107"  "1952"    "168439"  "819353"  "760967"
## [1093] "238205"  "186133"  "207126"  "2307779" "190639"  "195207"  "184394"
## [1100] "4459634" "179486"  "193191"  "13986"   "185034"  "2506486" "186748"
## [1107] "91557"   "336559"  "207994"  "4302904" "4453304" "91962"   "177792"
## [1114] "182674"  "186735"  "199524"  "688934"  "114462"  "4473129" "4405104"
## [1121] "122049"  "110060"  "125624"  "189334"  "4458758" "3569942" "797229"
## [1128] "3825935" "4362337" "198503"  "4356080" "4387246" "4479443" "378660"
## [1135] "4390330" "2331530" "4379889" "4386018" "293307"  "226338"  "4391009"
## [1142] "4334770" "73000"   "230232"  "344154"  "2500766" "4420570" "237081"
## [1149] "3997242" "1646259" "208790"  "299302"  "3462224" "4471279" "191355"
## [1156] "319197"  "368486"  "183147"  "3186216" "2407149" "606927"  "173965"
## [1163] "4481195" "2816372" "194110"  "592616"  "4459196" "207487"  "175373"
## [1170] "322835"  "4346374" "4404577" "363442"  "529979"  "4464445" "189067"
## [1177] "4373156" "309391"  "196526"  "190864"  "177421"  "2498994" "190991"
## [1184] "197581"  "189381"  "288362"  "185148"  "4381092" "4431803" "195548"
## [1191] "90487"   "1679707" "519746"  "1838560" "4408423" "510295"  "865469"
## [1198] "307680"  "1860111" "199421"  "4468891" "1763286" "1624383" "2063400"
## [1205] "193915"  "175520"  "4476065" "4029584" "3197419" "188077"  "289927"
## [1212] "2555599" "261590"  "109413"  "2272791" "177910"  "4454025" "3040024"
## [1219] "193831"  "430191"  "191999"  "4345285" "4343184" "191779"  "524848"
## [1226] "369635"  "3671383" "3186388" "195522"  "4380813" "180401"  "338177"
## [1233] "4294457" "356360"  "4449458" "4377715" "45296"   "4474951" "505670"
## [1240] "4337431" "185570"  "637546"  "4352349" "4367187" "144814"  "4481624"
## [1247] "4365141" "3359884" "743082"  "192342"  "203579"  "3441309" "119010"
## [1254] "67813"   "4310208" "4475758" "4473664" "366451"  "345540"  "1023075"
## [1261] "1115481" "1024529" "1130824" "746679"  "228556"  "370086"  "495451"
## [1268] "4409962" "1823053" "4469007" "628226"  "1888072" "4345397" "125270"
## [1275] "197499"  "583656"  "181161"  "192070"  "193679"  "174885"  "194320"
```

```
## [1282] "185864"  "1099710" "260414"  "174147"  "539581"  "188047"  "194219"
## [1289] "179319"  "3300804" "4393172" "4346675" "194001"  "370225"  "189524"
## [1296] "196982"  "188194"  "197343"  "195294"  "269386"  "197624"  "3697034"
## [1303] "4468892" "169364"  "181174"  "181008"  "166869"  "2768843" "174571"
## [1310] "110562"  "188707"  "2985051" "184770"  "4326869" "4477861" "173876"
## [1317] "197099"  "188454"  "199218"  "197203"  "194215"  "190171"  "843357"
## [1324] "192323"  "329703"  "193769"  "190649"  "185141"  "4472721" "186463"
## [1331] "2996838" "230421"  "187038"  "198190"  "4425495" "187868"  "198221"
## [1338] "175148"  "181176"  "300952"  "179785"  "197750"  "194559"  "1906635"
## [1345] "192818"  "1135084" "353615"  "182033"  "534957"  "2388088" "192370"
## [1352] "192839"  "2487129" "185428"  "3325758" "189176"  "328659"  "179663"
## [1359] "130663"  "197661"  "188333"  "332732"  "2170526" "538322"  "188851"
## [1366] "185861"  "2430693" "266274"  "191833"  "179729"  "1722623" "4386507"
## [1373] "176008"  "2200896" "134174"  "113069"  "1126851" "187689"  "199710"
## [1380] "2658061" "3474081" "189478"  "357305"  "4435655" "185583"  "297182"
## [1387] "4465472" "194258"  "186896"  "755148"  "521795"  "4429335" "495017"
## [1394] "25461"   "194761"  "181269"  "15728"   "494906"  "181853"  "1077373"
## [1401] "183824"  "179609"  "183210"  "4331723" "1100471" "189828"  "360238"
## [1408] "4476950" "3507744" "182653"  "180421"  "306066"  "181871"  "178965"
## [1415] "1646171" "186981"  "300374"  "232900"  "207340"  "553150"  "187404"
## [1422] "211720"  "178852"  "4421273" "186448"  "2256425" "3841096" "216599"
## [1429] "185164"  "564334"  "190913"  "1076818" "4349946" "360890"  "315223"
## [1436] "179847"  "178712"  "145856"  "718358"  "2442706" "839684"  "197890"
## [1443] "151870"  "186352"  "234443"  "4449427" "4466707" "4335484" "586453"
## [1450] "4468466" "194371"  "180107"  "4443846" "4370941" "157338"  "4472195"
## [1457] "4330001" "732765"  "345354"  "578511"  "2774254" "708285"  "189309"
## [1464] "214036"  "583500"  "191734"  "195494"  "186452"  "258099"  "199286"
## [1471] "2877100" "288138"  "175508"  "300378"  "1146771" "110192"  "188966"
## [1478] "3422630" "276484"  "148352"  "4364405" "172233"  "836693"  "3856408"
## [1485] "527988"  "563803"  "324015"  "341730"  "339566"  "2066056" "3390534"
## [1492] "193148"  "4460786" "193654"  "197052"  "174792"  "4456702" "3648884"
## [1499] "340711"  "181452"  "191476"  "3797933" "178708"  "4212012" "193129"
## [1506] "191222"  "561607"  "4364464" "113003"  "4475642" "2123719" "194036"
## [1513] "4387453" "177134"  "189588"  "198117"  "4404731" "54563"   "180136"
## [1520] "556240"  "216710"  "4435400" "178478"  "389371"  "2927689" "180473"
## [1527] "110144"  "227565"  "161762"  "2949328" "215231"  "204126"  "136127"
## [1534] "4435784" "183532"  "4347652" "184320"  "3141342" "4350477" "178183"
## [1541] "185059"  "368236"  "2391105" "193480"  "2403301" "4476561" "4466220"
## [1548] "188348"  "178557"  "210095"  "4466275" "352529"  "194372"  "2688035"
## [1555] "4472174" "562410"  "3195723" "330414"  "176862"  "197397"  "181827"
## [1562] "3692427" "293511"  "4474759" "213870"  "4455767" "4460021" "365167"
## [1569] "303326"  "814442"  "194380"  "193969"  "579244"  "195166"  "340759"
## [1576] "842786"  "182116"  "4354486" "2420187" "4450214" "157453"  "199534"
## [1583] "178018"  "2119695" "197698"  "197696"  "193852"  "184497"  "572889"
## [1590] "181059"  "175642"  "191551"  "175036"  "840376"  "4412540" "4450194"
## [1597] "2530636" "4442899" "4437359" "196513"  "2368865" "147702"  "4439398"
## [1604] "186772"  "309929"  "177207"  "1776714" "183093"  "197274"  "4403689"
## [1611] "839964"  "323135"  "193797"  "4391326" "288930"  "581201"  "192741"
## [1618] "4377149" "187225"  "189937"  "158625"  "191148"  "2532909" "185584"
## [1625] "4425669" "197761"  "178631"  "194734"  "4332878" "343257"  "332718"
## [1632] "4254313" "4405423" "4366391" "295485"  "323106"  "188315"  "196918"
## [1639] "332450"  "801210"  "173927"  "180090"  "189760"  "180257"  "174304"
## [1646] "4456027" "300662"  "4468805" "288004"  "304973"  "360311"  "4403113"
## [1653] "176381"  "181217"  "4451907" "178082"  "317814"  "575768"  "198626"
```

```
## [1660] "198127"  "194995"  "179291"  "189624"  "194339"  "4375889" "3544569"
## [1667] "194443"  "161007"  "197708"  "4484075" "177062"  "198866"  "190679"
## [1674] "308081"  "180842"  "193987"  "322571"  "194868"  "4360055" "191391"
## [1681] "4379646" "183224"  "4226619" "194727"  "291648"  "287875"  "190169"
## [1688] "522433"  "130763"  "175180"  "358781"  "192462"  "180572"  "363646"
## [1695] "183457"  "185575"  "183054"  "319275"  "190987"  "583134"  "184114"
## [1702] "3530697" "194297"  "185836"  "173851"  "194584"  "4409280" "4366843"
## [1709] "362793"  "330469"  "180771"  "158649"  "4366089" "146564"  "4336940"
## [1716] "185814"  "194471"  "194324"  "176469"  "307984"  "179572"  "3807411"
## [1723] "208901"  "199256"  "192438"  "186011"  "201772"  "3235048" "188449"
## [1730] "181074"  "335670"  "187504"  "178331"  "326390"  "2137906" "198583"
## [1737] "216010"  "196724"  "2730944" "4458634" "1614788" "187468"  "552235"
## [1744] "197460"  "193778"  "187924"  "364227"  "188262"  "186471"  "4463532"
## [1751] "187180"  "4390972" "186906"  "184980"  "341920"  "4232048" "180215"
## [1758] "182036"  "175836"  "4443094" "327218"  "4469032" "166848"  "232828"
## [1765] "192437"  "3302038" "189877"  "305288"  "199344"  "322505"  "3302039"
## [1772] "182431"  "179273"  "851733"  "196017"  "302160"  "1047077" "215097"
## [1779] "924547"  "4290144" "581021"  "4359797" "197286"  "4227110" "4297420"
## [1786] "928538"  "1132942" "281015"  "772515"  "1010113" "4376230" "4384058"
## [1793] "167950"  "4453060" "1995363" "691423"  "348680"  "839152"  "305561"
## [1800] "4454257" "182483"  "4318284" "4363066" "1811927" "4387092" "289925"
## [1807] "4405869" "196391"  "4396656" "1105376" "4041792" "300297"  "237444"
## [1814] "191783"  "196176"  "248126"  "303161"  "287813"  "351020"  "4352805"
## [1821] "134671"  "4328189" "189292"  "174499"  "4352747" "194925"  "184729"
## [1828] "368175"  "342666"  "4477719" "174738"  "4428544" "174360"  "179576"
## [1835] "544419"  "212686"  "325850"  "4433417" "49088"   "305224"  "178242"
## [1842] "352852"  "157054"  "577710"  "189793"  "175932"  "195947"  "182933"
## [1849] "4070490" "342231"  "176157"  "193061"  "1986324" "4411138" "552376"
## [1856] "184036"  "258375"  "175650"  "289452"  "156357"  "136518"  "348009"
## [1863] "187384"  "4469233" "688923"  "4442508" "182735"  "176104"  "4420417"
## [1870] "4358921" "688800"  "336063"  "179861"  "192592"  "157772"  "4319899"
## [1877] "4335578" "2229500" "110625"  "4406616" "4450415" "3598692" "196831"
## [1884] "193829"  "192696"  "4429239" "179436"  "4448928" "4351284" "173810"
## [1891] "296420"  "4475469" "849642"  "4315658" "71165"   "3610691" "4468384"
## [1898] "518820"  "4437746" "198468"  "195029"  "4352494" "4349261" "1667433"
## [1905] "4427290" "315958"  "1107755" "178849"  "348863"  "174045"  "4434268"
## [1912] "331150"  "4304475" "181826"  "321517"  "2800178" "4415649" "4459414"
## [1919] "182588"  "4374853" "695909"  "1124877" "1959881" "3889756" "528628"
## [1926] "4418165" "242298"  "3680006" "174960"  "12562"   "147195"  "64384"
## [1933] "4102199" "199555"  "308444"  "3270612" "193975"  "181047"  "519763"
## [1940] "372146"  "352747"  "178994"  "177828"  "891034"  "311402"  "165118"
## [1947] "187917"  "181718"  "333114"  "195358"  "287700"  "338301"  "191250"
## [1954] "186381"  "179715"  "196581"  "304757"  "4306587" "4424063" "306412"
## [1961] "4405482" "4437748" "3096648" "361702"  "196381"  "128382"  "103166"
## [1968] "4428313" "4397402" "1108437" "137026"  "4354103" "12564"   "275237"
## [1975] "198928"  "4393073" "3221787" "294794"  "363731"  "214031"  "4415390"
## [1982] "363997"  "188044"  "4072210" "225088"  "4305372" "350503"  "4202174"
## [1989] "180974"  "4332078" "4428676" "129692"  "290241"  "158722"  "230479"
## [1996] "190863"  "175336"  "3696710" "2182669" "146086"  "192255"  "4326084"
## [2003] "308157"  "181466"  "192654"  "193314"  "180857"  "221429"  "198248"
## [2010] "189936"  "2978122" "186272"  "2399341" "1986406" "102471"  "3693042"
## [2017] "759816"  "294254"  "19611"   "4429608" "4397439" "179018"  "4451906"
## [2024] "842193"  "4355379" "191718"  "4297222" "319411"  "4004998" "1068836"
## [2031] "4371463" "168071"  "357930"  "189754"  "175967"  "303269"  "129401"
```

```
## [2038] "3278416" "183925"  "183804"  "348642"  "186851"  "113581"  "1657491"
## [2045] "309107"  "3438642" "293594"  "4396298" "4153054" "190961"  "185802"
## [2052] "4336943" "1111522" "3579831" "176980"  "1012948" "179744"  "847934"
## [2059] "3806695" "4425663" "183439"  "4483045" "540055"  "239863"  "179435"
## [2066] "364563"  "329820"  "4422405" "2932557" "181167"  "584176"  "232696"
## [2073] "317286"  "3228684" "556126"  "3101394" "1147925" "4480176" "186358"
## [2080] "190241"  "179392"  "178713"  "326936"  "298937"  "2883968" "4453773"
## [2087] "4472551" "136526"  "320888"  "193868"  "4407747" "178970"  "815422"
## [2094] "1034960" "849393"  "522595"  "158113"  "544996"  "4307265" "196590"
## [2101] "4473883" "564290"  "386273"  "178253"  "4178726" "185914"  "366722"
## [2108] "174943"  "179989"  "305946"  "1028036" "4480359" "181961"  "4154872"
## [2115] "548503"  "4308811" "177301"  "4357713" "302333"  "2699610" "2829179"
## [2122] "4381422" "4416614" "663500"  "178064"  "547913"  "3465233" "2024346"
## [2129] "185972"  "195723"  "710275"  "178809"  "4423553" "190772"  "195081"
## [2136] "177758"  "185603"  "179402"  "182416"  "302321"  "4459940" "363430"
## [2143] "186233"  "351163"  "327851"  "174019"  "368412"  "158027"  "4404507"
## [2150] "130048"  "3409355" "4368575" "815573"  "840961"  "358712"  "4358599"
## [2157] "685156"  "3550973" "179662"  "3609803" "316728"  "2232355" "14920"
## [2164] "2134134" "333768"  "579541"  "292758"  "4333921" "2820255" "353330"
## [2171] "178773"  "4474760" "361864"  "4397098" "182142"  "288258"  "610111"
## [2178] "105409"  "725198"  "4012414" "302880"  "184314"  "268506"  "960871"
## [2185] "4480861" "4415965" "724147"  "646800"  "4338624" "107044"  "4307122"
## [2192] "304088"  "195871"  "214682"  "1106254" "347783"  "205613"  "846798"
## [2199] "177518"  "158423"  "16340"   "414949"  "4310326" "4353913" "180982"
## [2206] "1145262" "747563"  "157470"  "1131759" "2423305" "4300690" "342844"
## [2213] "296858"  "279048"  "511229"  "300618"  "109633"  "4465124" "308386"
## [2220] "191278"  "1148151" "113542"  "158215"  "2377731" "41128"   "4352875"
## [2227] "320915"  "3785400" "574720"  "2965498" "4399109" "4299126" "185593"
## [2234] "158270"  "342777"  "842596"  "144395"  "229348"  "187121"  "208972"
```
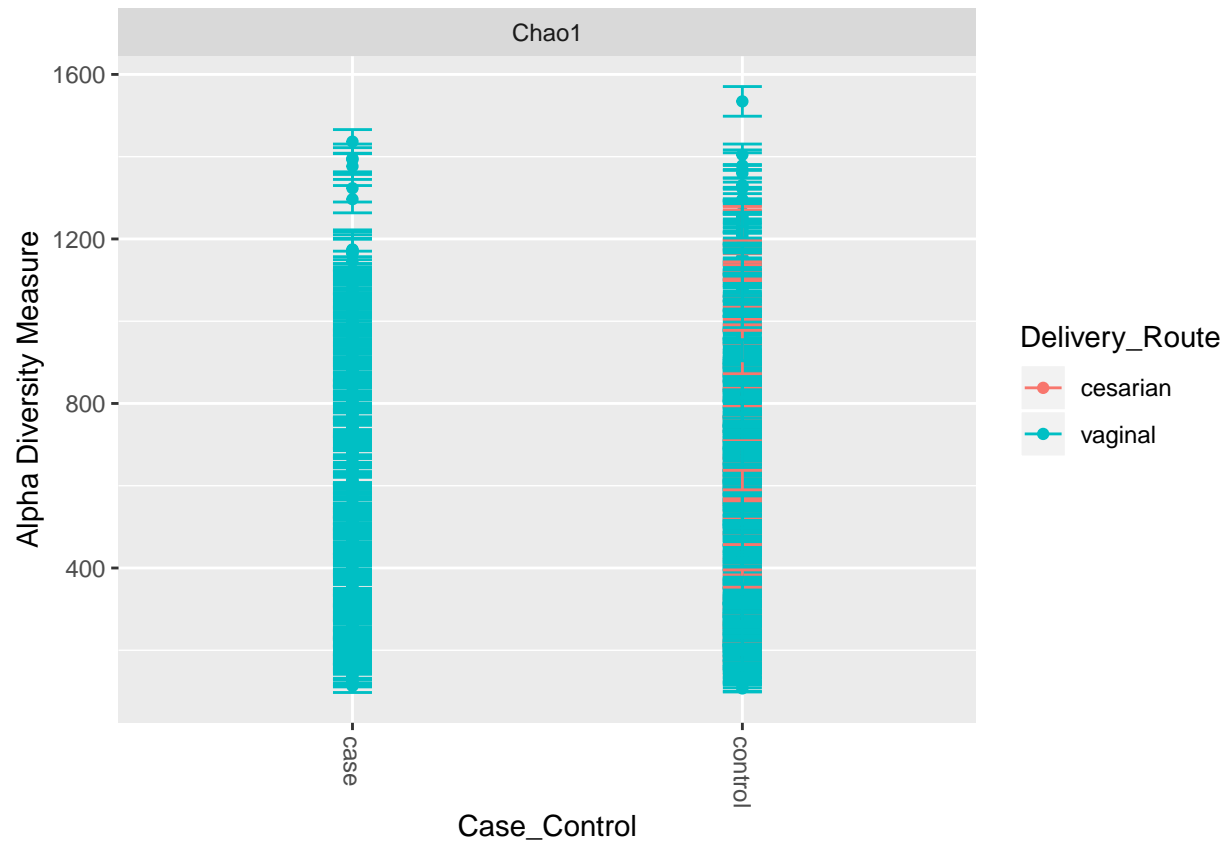
```
sample_variables(my_phyloseq_object)
```

```
## [1] "Subject_ID"       "Case_Control"     "Gender"
## [4] "Delivery_Route"   "Age_at_Collection"
```
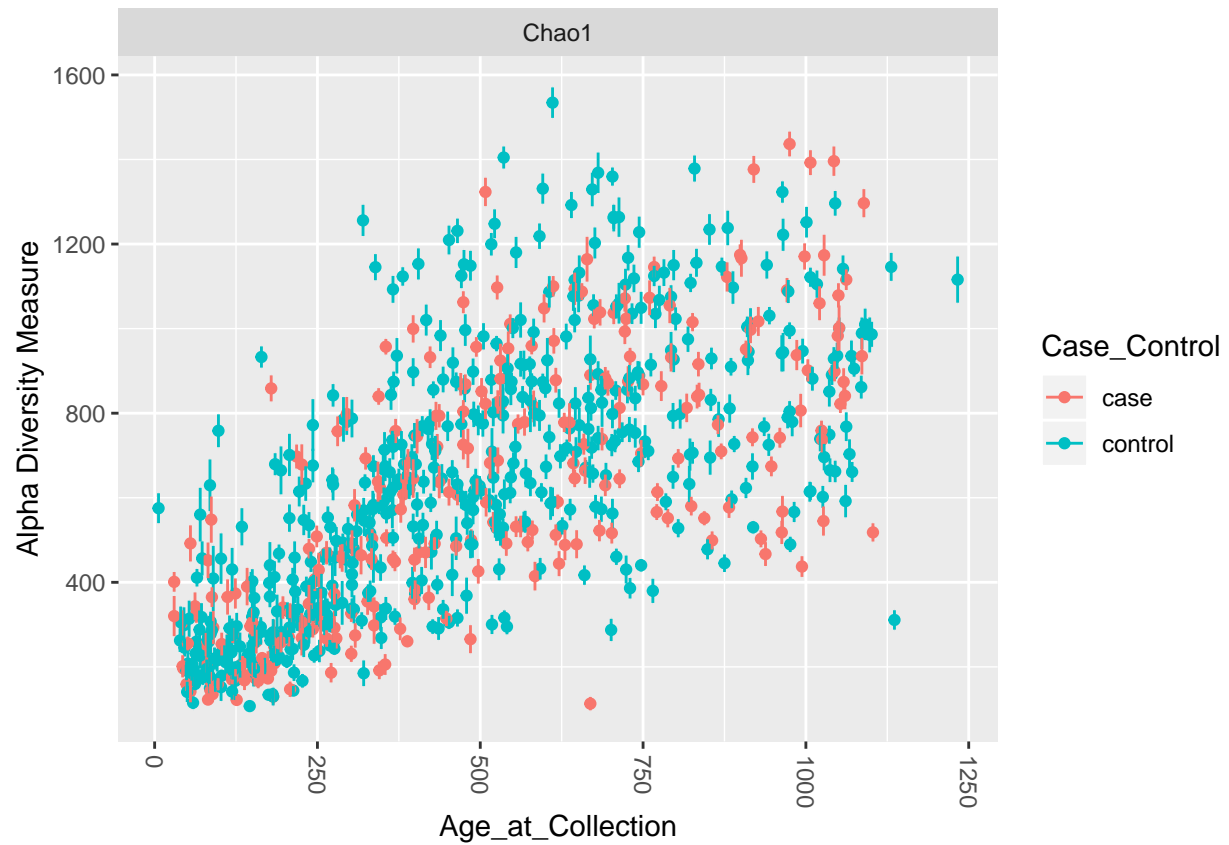
Generate alpha diversity and ordination plots

```
#Alpha Diversity
#Chao1:non parametric method for estimating the number of species in a community
#shannon:places greater weight on richness.
plot_richness(my_phyloseq_object,x = "Case_Control", color="Delivery_Route", measures= "Chao1")
```
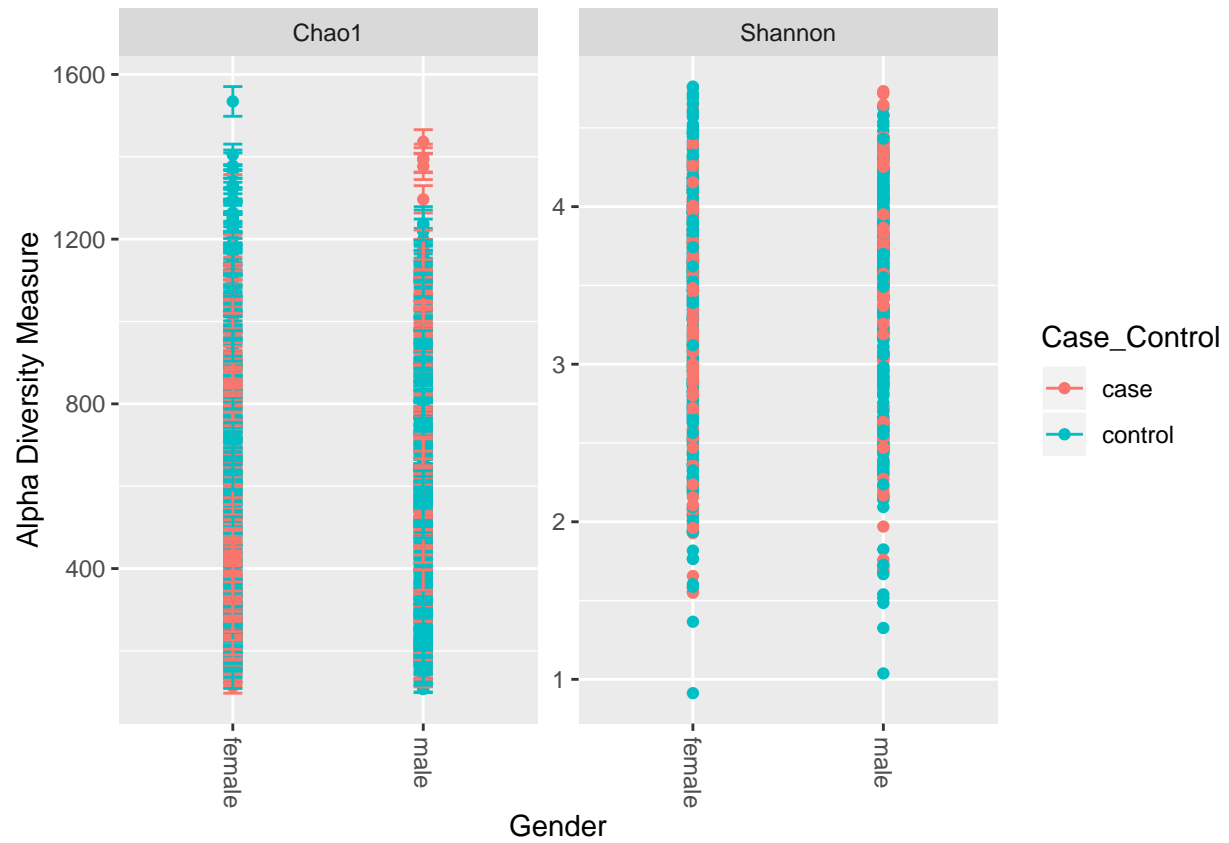
```
plot_richness(my_phyloseq_object,x = "Age_at_Collection", color="Case_Control", measures= "Chao1")
```
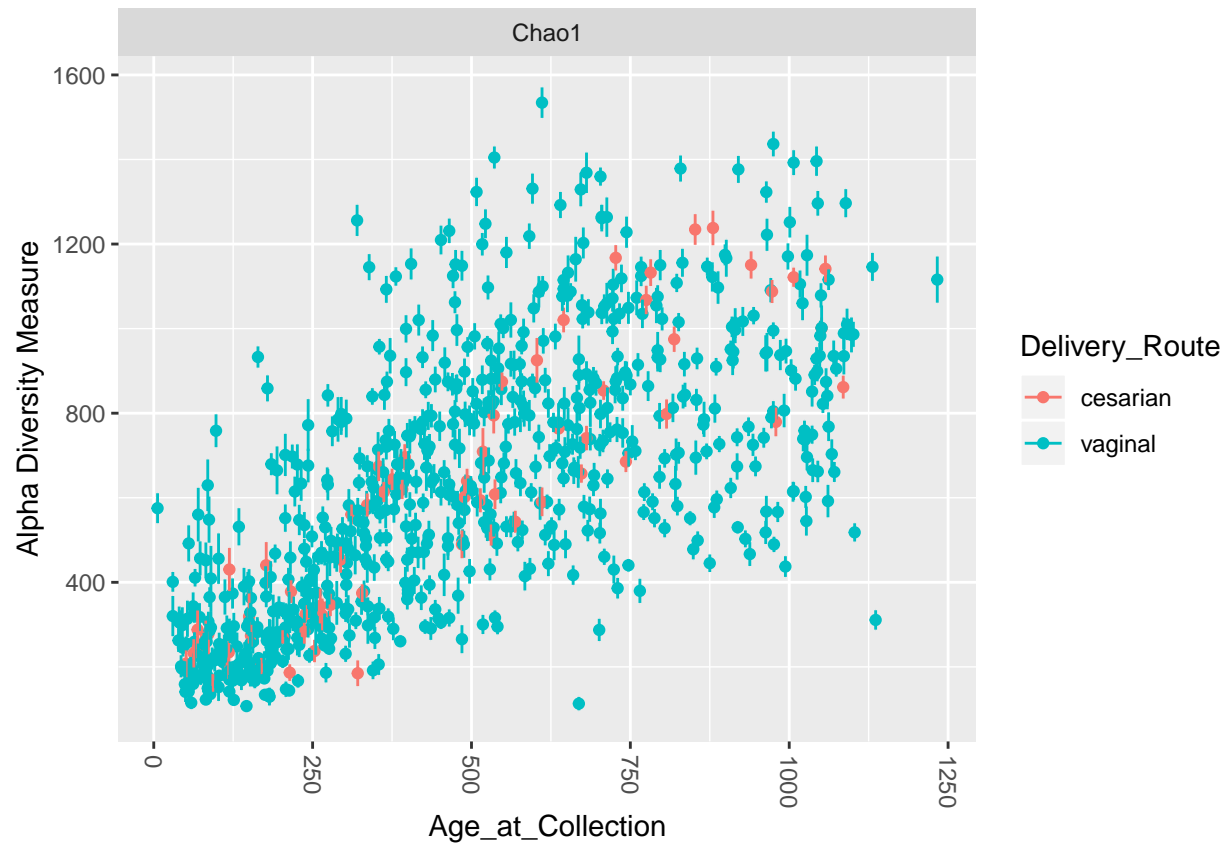
```
plot_richness(my_phyloseq_object,x = "Gender", color="Case_Control", measures= c("Chao1", "shannon"))
```
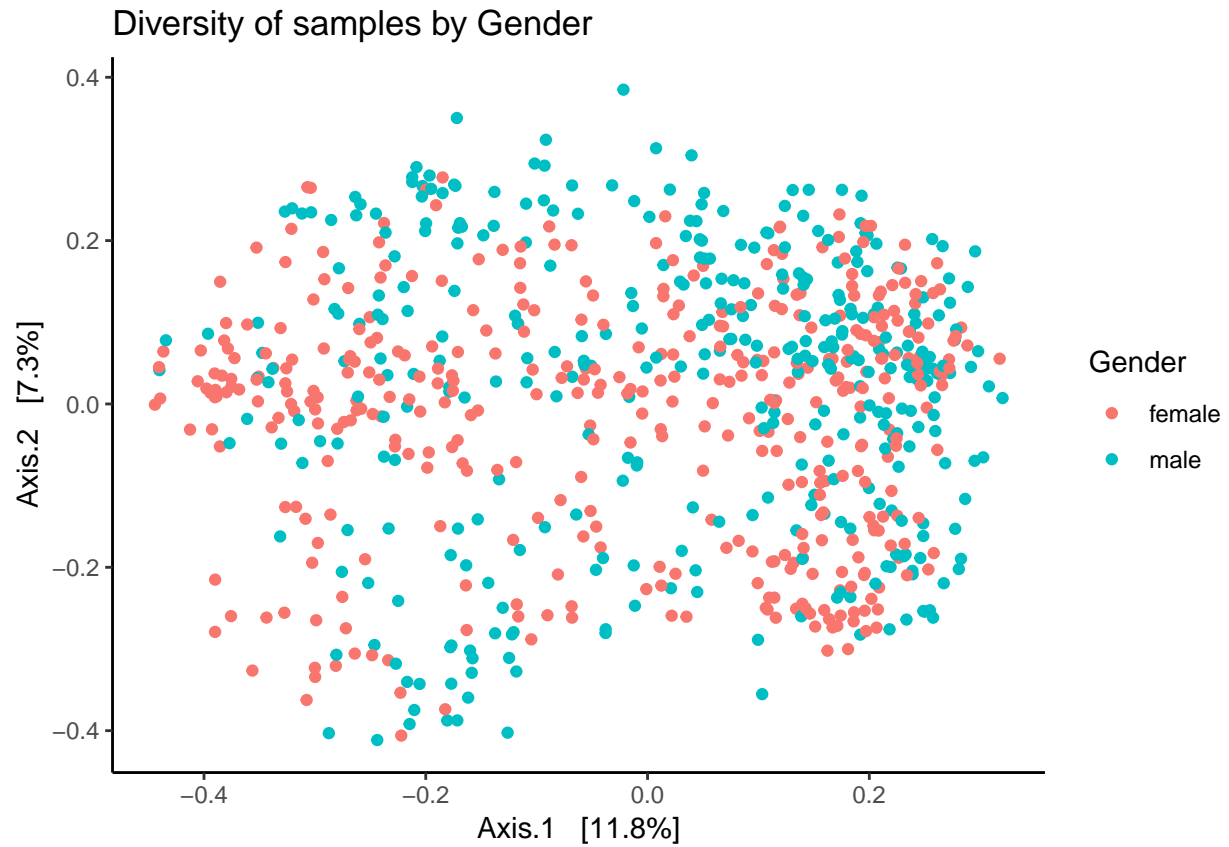
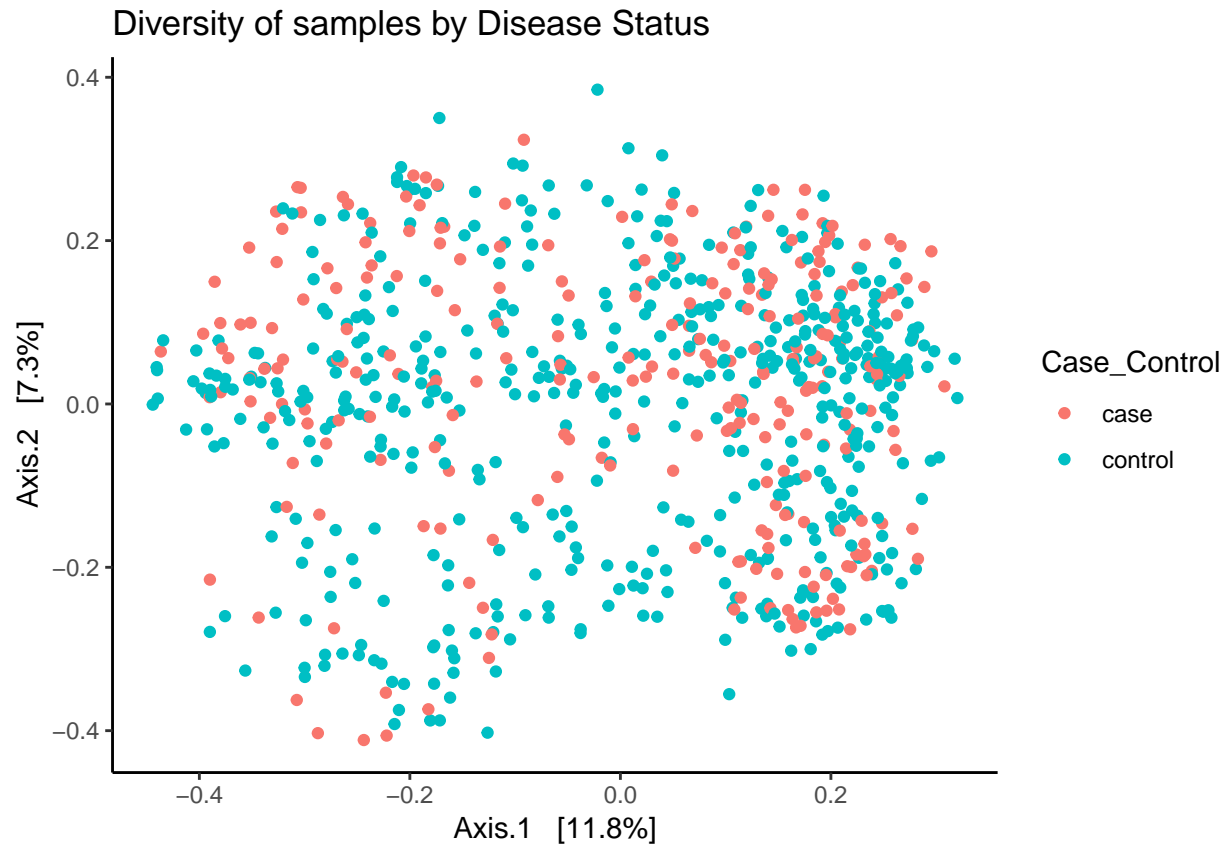## Warning: Removed 777 rows containing missing values (geom_errorbar).

```
plot_richness(my_phyloseq_object,x = "Age_at_Collection", color="Delivery_Route", measures= "Chao1")
```
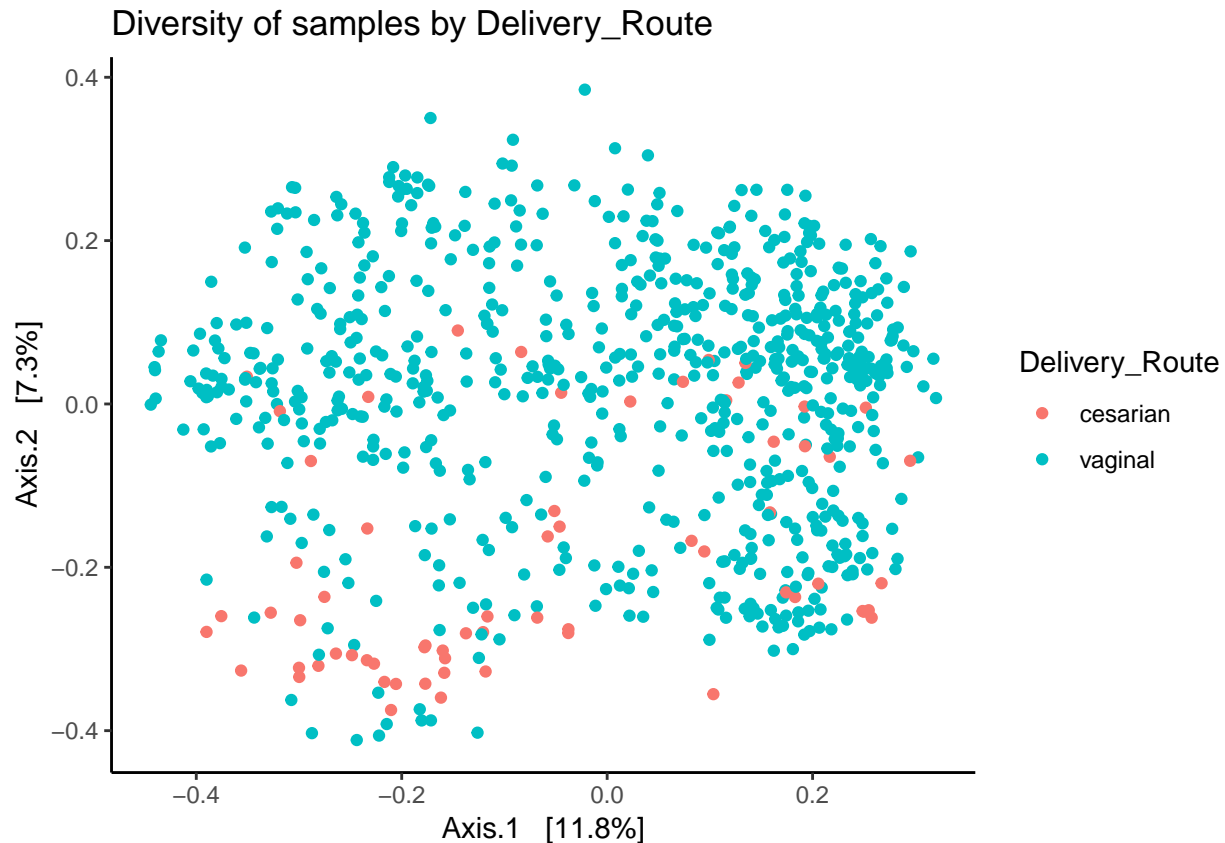
```
#Ordination Plots
ordinate(my_phyloseq_object, "PCoA", "bray")%>%
  plot_ordination(my_phyloseq_object, .,color = "Gender", title = "Diversity of samples by Gender")+
  theme_classic()
```

Diversity of samples by Gender


```
ordinate <- ordinate(my_phyloseq_object, "PCoA", "bray")
plot_ordination(my_phyloseq_object, ordinate,color = "Case_Control", title = "Diversity of samples by D:
  theme_classic()
```

Diversity of samples by Disease Status


```
ordinate <- ordinate(my_phyloseq_object, "PCoA", "bray")
plot_ordination(my_phyloseq_object, ordinate,color = "Delivery_Route", title = "Diversity of samples by
  theme_classic()
```

Diversity of samples by Delivery_Route

Interpretation and examination for any observed patterns.

- There was no diversity in the cases, all cases were born through the vaginal delivery route. None of the participants born through caeserian section developed the disease.

- Cases and controls were almost richly and evenly distributed across all the values in age at collection. All participants regardless of their age could be either cases or controls.

- Cases and Controls were almost evenly and richly distributed in both males and females but both cases and controls were more in females as compared to males.

4:Perform a differential abundance using Deseq2

```
library(DESeq2) #call package DEseq2
```

```
## Loading required package: S4Vectors
```

```
## Loading required package: stats4
```

```
## Loading required package: BiocGenerics
```

```
## Loading required package: parallel
```

```
##
## Attaching package: 'BiocGenerics'
```

```
## The following objects are masked from 'package:parallel':
##
##     clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##     clusterExport, clusterMap, parApply, parCapply, parLapply,
##     parLapplyLB, parRapply, parSapply, parSapplyLB

## The following objects are masked from 'package:dplyr':
##
##     combine, intersect, setdiff, union

## The following objects are masked from 'package:stats':
##
##     IQR, mad, sd, var, xtabs

## The following objects are masked from 'package:base':
##
##     anyDuplicated, append, as.data.frame, basename, cbind, colMeans,
##     colnames, colSums, dirname, do.call, duplicated, eval, evalq,
##     Filter, Find, get, grep, grepl, intersect, is.unsorted, lapply,
##     lengths, Map, mapply, match, mget, order, paste, pmax, pmax.int,
##     pmin, pmin.int, Position, rank, rbind, Reduce, rowMeans, rownames,
##     rowSums, sapply, setdiff, sort, table, tapply, union, unique,
##     unsplit, which, which.max, which.min

##
## Attaching package: 'S4Vectors'

## The following objects are masked from 'package:dplyr':
##
##     first, rename

## The following object is masked from 'package:tidyr':
##
##     expand

## The following object is masked from 'package:base':
##
##     expand.grid

## Loading required package: IRanges

##
## Attaching package: 'IRanges'

## The following object is masked from 'package:phyloseq':
##
##     distance

## The following objects are masked from 'package:dplyr':
##
##     collapse, desc, slice

## The following object is masked from 'package:purrr':
##
##     reduce

## Loading required package: GenomicRanges

## Loading required package: GenomeInfoDb

## Loading required package: SummarizedExperiment
```

```
## Loading required package: Biobase

## Welcome to Bioconductor
##
##     Vignettes contain introductory material; view with
##     'browseVignettes()'. To cite Bioconductor, see
##     'citation("Biobase")', and for packages 'citation("pkgname")'.

##
## Attaching package: 'Biobase'

## The following object is masked from 'package:phyloseq':
##
##     sampleNames

## Loading required package: DelayedArray

## Loading required package: matrixStats

##
## Attaching package: 'matrixStats'

## The following objects are masked from 'package:Biobase':
##
##     anyMissing, rowMedians

## The following object is masked from 'package:dplyr':
##
##     count

## Loading required package: BiocParallel

##
## Attaching package: 'DelayedArray'

## The following objects are masked from 'package:matrixStats':
##
##     colMaxs, colMins, colRanges, rowMaxs, rowMins, rowRanges

## The following object is masked from 'package:purrr':
##
##     simplify

## The following objects are masked from 'package:base':
##
##     aperm, apply
```

```r
#DESeq: Creating a DESeq object
#add 1 because some of the values in the OTU table were 0, Deseq can not work with such data.
#convert Phyloseq object from phyloseq format to deseq dataset and estimate dispersions
my_phyloseq_object2 <- phyloseq(otu_table(OTU_table_matrix + 1,taxa_are_rows=TRUE),tax_table(taxonomymat
sample_data(my_phyloseq_object2) <- phyloseq_diabetes_data
sample_data(my_phyloseq_object2)["Case_Control"] <- relevel(unlist(sample_data(my_phyloseq_object2)["Ca
my_casecontrol = phyloseq_to_deseq2(my_phyloseq_object2, ~ Case_Control)
```

```
## converting counts to integer mode
```

```r
#DESeq test:tests for dispersions
my_casecontrol = DESeq(my_casecontrol, test="Wald", fitType="parametric")
```

```
## estimating size factors

## estimating dispersions
```
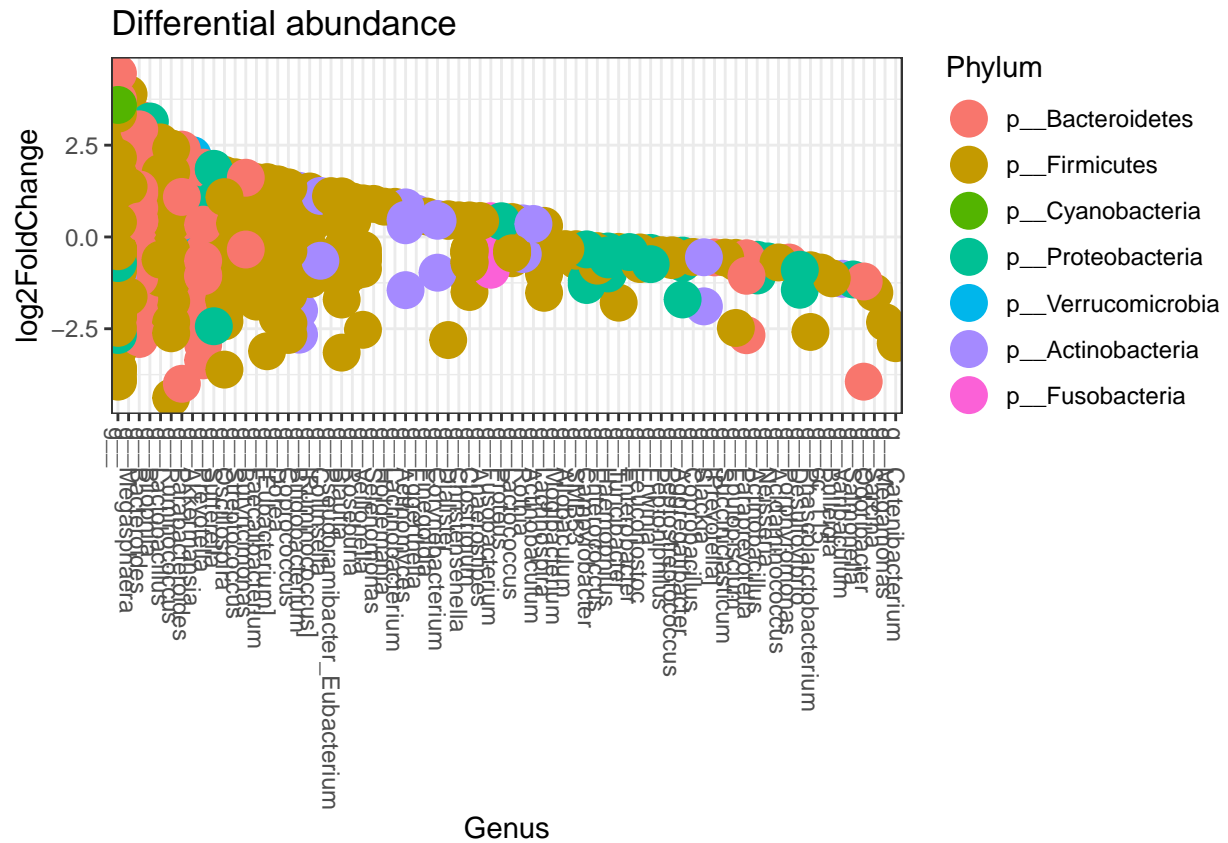
```
## gene-wise dispersion estimates

## mean-dispersion relationship

## -- note: fitType='parametric', but the dispersion trend was not well captured by the
##    function: y = a/x + b, and a local regression fit was automatically substituted.
##    specify fitType='local' or 'mean' to avoid this message next time.

## final dispersion estimates

## fitting model and testing

## -- replacing outliers and refitting for 991 genes
## -- DESeq argument 'minReplicatesForReplace' = 7
## -- original counts are preserved in counts(dds)

## estimating dispersions

## fitting model and testing
```

```r
#Results table:results function creates a table of results.
my_deseq_results = results(my_casecontrol, cooksCutoff = FALSE)
alpha = 0.01
my_casetab = my_deseq_results[which(my_deseq_results$padj < alpha), ]
my_casetab = cbind(as(my_casetab, "data.frame"), as(tax_table(my_phyloseq_object2)[rownames(my_casetab)

#plot to show OTUS that are significantly different.
theme_set(theme_bw())
scale_fill_discrete <- function(palname = "Set1", ...) {
  scale_fill_brewer(palette = palname, ...)
}
# Phylum order
x = tapply(my_casetab$log2FoldChange, my_casetab$Phylum, function(x) max(x))
x = sort(x, TRUE)
my_casetab$Phylum = factor(as.character(my_casetab$Phylum), levels=names(x))
# Genus order
x = tapply(my_casetab$log2FoldChange, my_casetab$Genus, function(x) max(x))
x = sort(x, TRUE)
my_casetab$Genus = factor(as.character(my_casetab$Genus), levels=names(x))
ggplot(my_casetab, aes(x=Genus, y=log2FoldChange, color=Phylum)) + geom_point(size=6) +
  theme(axis.text.x = element_text(angle = -90, hjust = 0, vjust=0.5)) + ggtitle("Differential abundance
```

Differential abundance

Interpretation of differential abundance

As seen in the graph above, Phylum Firmicutes is the most abundant of them all and phylum Fusobacteria is the least abundant.