

```

#Intro to Data Science - Final Project - Part 1
#East Section
#March 30, 2018

#Group Members:
# Catherine Aslinger
# Sharif Nijim
# Gina O'Riordan

# initialize the environment
rm(list=ls())
#Load Libraries
library(tools)
library(tidyverse)

## Warning: package 'tidyverse' was built under R version 3.4.2

## — Attaching packages ————— tidyverse 1.2.1 —
— 

## ✓ ggplot2 2.2.1      ✓ purrr  0.2.4
## ✓ tibble  1.4.2      ✓ dplyr   0.7.4
## ✓ tidyr   0.7.2      ✓ stringr 1.2.0
## ✓ readr   1.1.1      ✓ forcats 0.2.0

## Warning: package 'tidyverse' was built under R version 3.4.2
## Warning: package 'tidyverse' was built under R version 3.4.2
## Warning: package 'tidyverse' was built under R version 3.4.2
## Warning: package 'tidyverse' was built under R version 3.4.2

## — Conflicts ————— tidyverse_conflicts() —
— 

## ✘ dplyr::filter() masks stats::filter()
## ✘ dplyr::lag()   masks stats::lag()

library(lubridate)

## Warning: package 'lubridate' was built under R version 3.4.2

##
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':
##
##     date

library(mosaic)

## Warning: package 'mosaic' was built under R version 3.4.3

## Loading required package: lattice

```

```
## Loading required package: ggformula
## Warning: package 'ggformula' was built under R version 3.4.3
##
## New to ggformula? Try the tutorials:
##   learnr::run_tutorial("introduction", package = "ggformula")
##   learnr::run_tutorial("refining", package = "ggformula")

## Loading required package: mosaicData

## Loading required package: Matrix

## Warning: package 'Matrix' was built under R version 3.4.2
##
## Attaching package: 'Matrix'

## The following object is masked from 'package:tidyverse':
##   expand
##
## The 'mosaic' package masks several functions from core packages in order to add
## additional features. The original behavior of these functions should not be affected by this.
##
## Note: If you use the Matrix package, be sure to load it BEFORE loading mosaic.

##
## Attaching package: 'mosaic'

## The following object is masked from 'package:Matrix':
##   mean
##
## The following objects are masked from 'package:dplyr':
##   count, do, tally
##
## The following object is masked from 'package:purrr':
##   cross
##
## The following objects are masked from 'package:stats':
##   binom.test, cor, cor.test, cov, fivenum, IQR, median,
##   prop.test, quantile, sd, t.test, var
```

```

## The following objects are masked from 'package:base':
##
##     max, mean, min, prod, range, sample, sum

library(rpart)

## Warning: package 'rpart' was built under R version 3.4.3

library(partykit)

## Warning: package 'partykit' was built under R version 3.4.3

## Loading required package: grid

## Loading required package: libcoin

## Warning: package 'libcoin' was built under R version 3.4.3

## Loading required package: mvtnorm

## Warning: package 'mvtnorm' was built under R version 3.4.3

library(randomForest)

## randomForest 4.6-12

## Type rfNews() to see new features/changes/bug fixes.

##
## Attaching package: 'randomForest'

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

## The following object is masked from 'package:ggplot2':
##
##     margin

library(class)
library(neuralnet)

##
## Attaching package: 'neuralnet'

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

#1 - Download the data from NeXus: FNA_cancer.csv
cancer <- read.csv("FNA_cancer.csv")
glimpse(cancer)

```

```

## Observations: 569
## Variables: 33
## $ id                               <int> 842302, 842517, 84300903, 84348301, 84...
## $ diagnosis                         <fct> M, ...
## $ radius_mean                       <dbl> 17.990, 20.570, 19.690, 11.420, 20.290...
## $ texture_mean                      <dbl> 10.38, 17.77, 21.25, 20.38, 14.34, 15....
## $ perimeter_mean                    <dbl> 122.80, 132.90, 130.00, 77.58, 135.10, ...
## $ area_mean                          <dbl> 1001.0, 1326.0, 1203.0, 386.1, 1297.0, ...
## $ smoothness_mean                   <dbl> 0.11840, 0.08474, 0.10960, 0.14250, 0....
## $ compactness_mean                  <dbl> 0.27760, 0.07864, 0.15990, 0.28390, 0....
## $ concavity_mean                   <dbl> 0.30010, 0.08690, 0.19740, 0.24140, 0....
## $ concave.points_mean              <dbl> 0.14710, 0.07017, 0.12790, 0.10520, 0....
## $ symmetry_mean                     <dbl> 0.2419, 0.1812, 0.2069, 0.2597, 0.1809...
## $ fractal_dimension_mean           <dbl> 0.07871, 0.05667, 0.05999, 0.09744, 0....
## $ radius_se                          <dbl> 1.0950, 0.5435, 0.7456, 0.4956, 0.7572...
## $ texture_se                         <dbl> 0.9053, 0.7339, 0.7869, 1.1560, 0.7813...
## $ perimeter_se                      <dbl> 8.589, 3.398, 4.585, 3.445, 5.438, 2.2...
## $ area_se                            <dbl> 153.40, 74.08, 94.03, 27.23, 94.44, 27...
## $ smoothness_se                      <dbl> 0.006399, 0.005225, 0.006150, 0.009110...
## $ compactness_se                     <dbl> 0.049040, 0.013080, 0.040060, 0.074580...
## $ concavity_se                       <dbl> 0.05373, 0.01860, 0.03832, 0.05661, 0....
## $ concave.points_se                 <dbl> 0.015870, 0.013400, 0.020580, 0.018670...
## $ symmetry_se                        <dbl> 0.03003, 0.01389, 0.02250, 0.05963, 0....
## $ fractal_dimension_se              <dbl> 0.006193, 0.003532, 0.004571, 0.009208...
## $ radius_worst                       <dbl> 25.38, 24.99, 23.57, 14.91, 22.54, 15....
## $ texture_worst                      <dbl> 17.33, 23.41, 25.53, 26.50, 16.67, 23...
## $ perimeter_worst                   <dbl> 184.60, 158.80, 152.50, 98.87, 152.20, ...
## $ area_worst                          <dbl> 2019.0, 1956.0, 1709.0, 567.7, 1575.0, ...
## $ smoothness_worst                   <dbl> 0.1622, 0.1238, 0.1444, 0.2098, 0.1374...
## $ compactness_worst                  <dbl> 0.6656, 0.1866, 0.4245, 0.8663, 0.2050...
## $ concavity_worst                   <dbl> 0.71190, 0.24160, 0.45040, 0.68690, 0....
## $ concave.points_worst              <dbl> 0.26540, 0.18600, 0.24300, 0.25750, 0....
## $ symmetry_worst                     <dbl> 0.4601, 0.2750, 0.3613, 0.6638, 0.2364...
## $ fractal_dimension_worst           <dbl> 0.11890, 0.08902, 0.08758, 0.17300, 0....
## $ X                                 <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA...

```

#2 - Perform basic exploratory data analysis

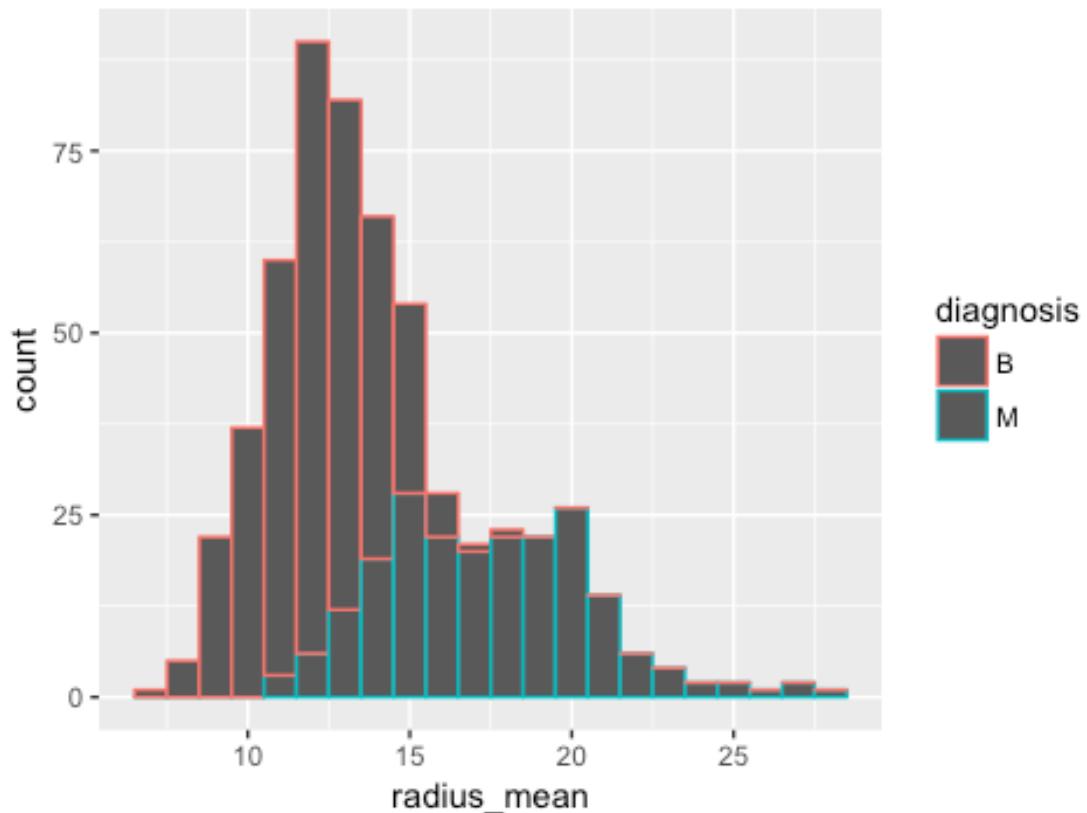
```

## Univariate EDA

#HISTOGRAMS
##Histograms of mean variables
#radius_mean
hist_radius_mean <- ggplot(data = cancer, aes(x = radius_mean)) + aes(color=diagnosis) +
  geom_histogram(binwidth = 1) + labs(title = "Histogram Radius Mean")
hist_radius_mean

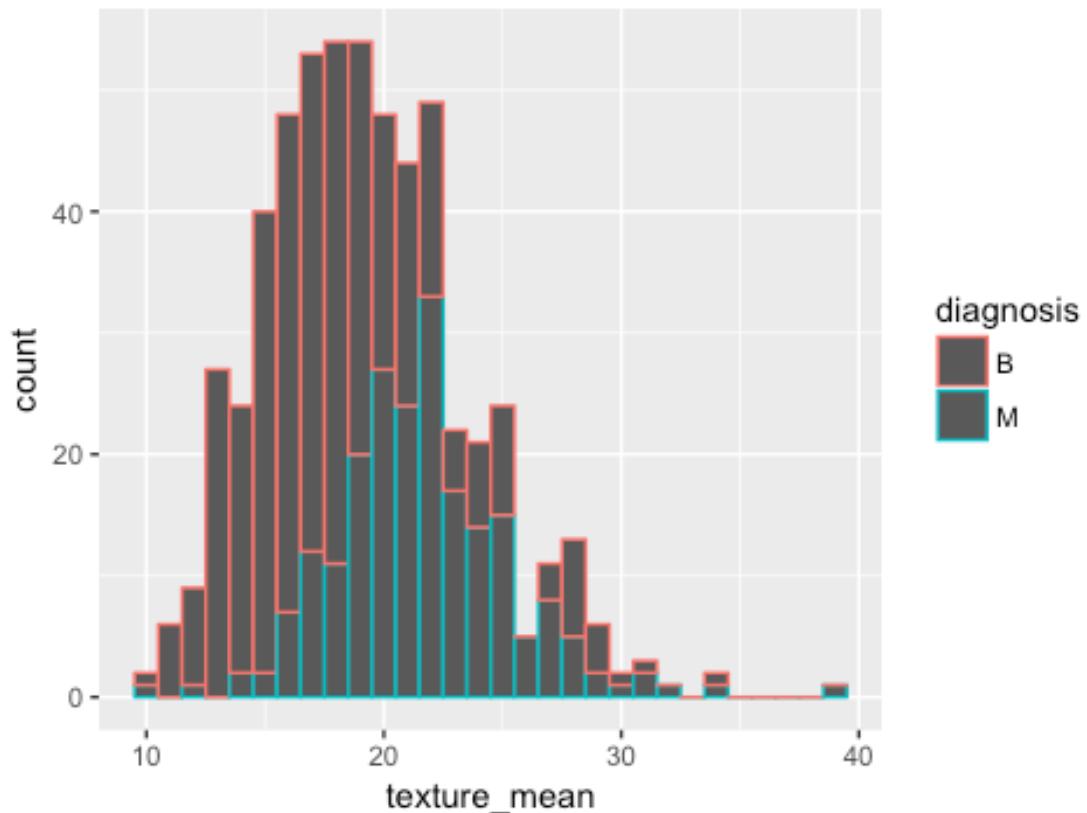
```

Histogram Radius Mean



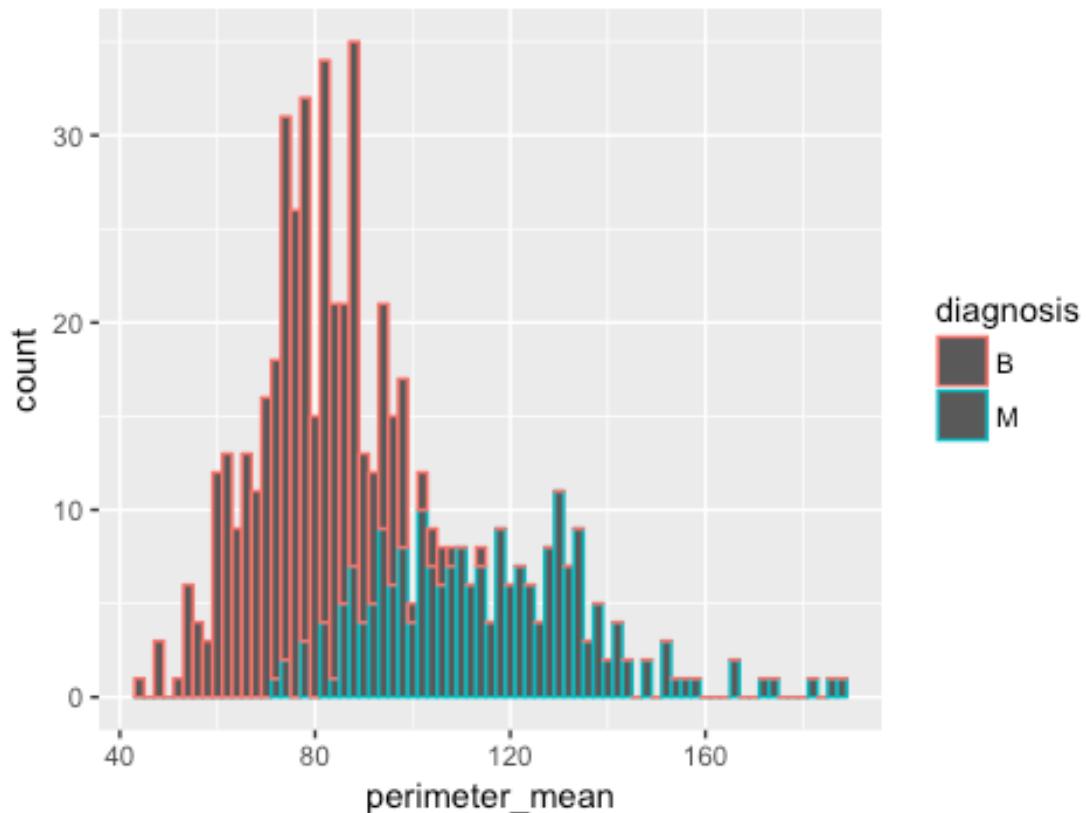
```
#texture_mean
hist_texture_mean <- ggplot(data = cancer, aes(x = texture_mean)) + aes(color = diagnosis) +
  geom_histogram(binwidth = 1) + labs(title = "Histogram Texture Mean")
hist_texture_mean
```

Histogram Texture Mean



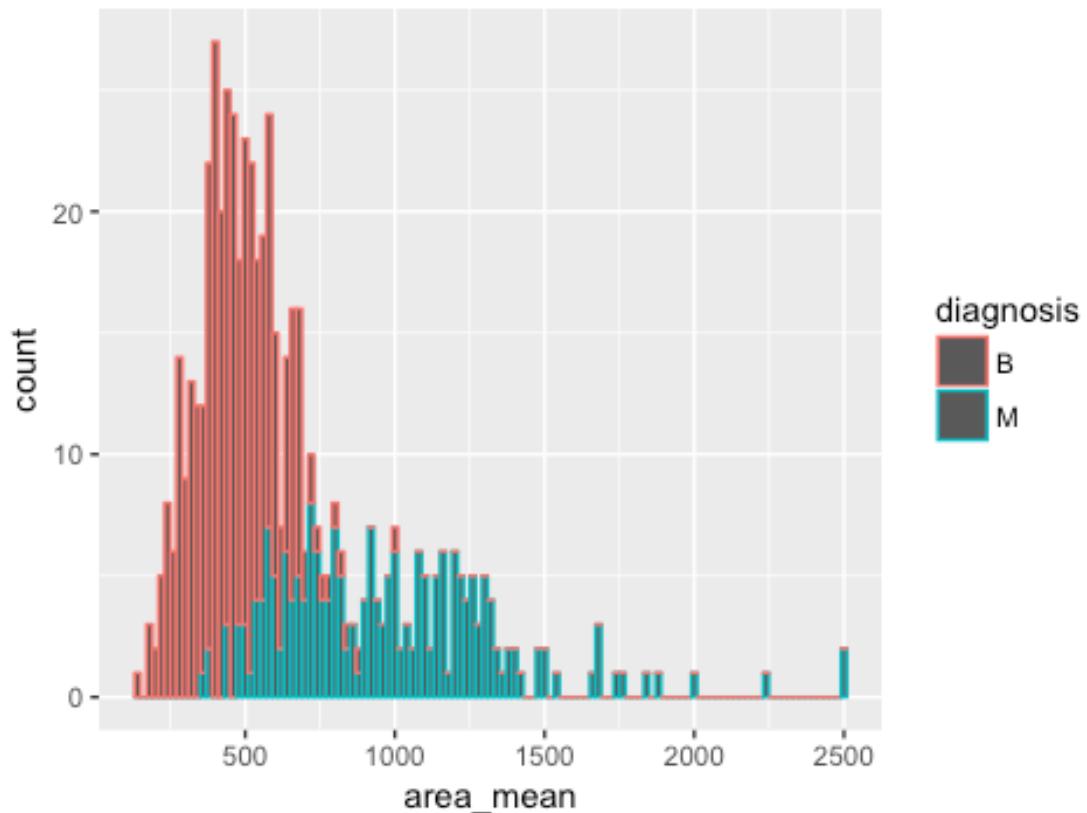
```
#perimeter_mean
hist_perimeter_mean <- ggplot(data = cancer, aes(x = perimeter_mean)) + aes(color=diagnosis) +
  geom_histogram(binwidth = 2) + labs(title = "Histogram Perimeter Mean")
hist_perimeter_mean
```

Histogram Perimeter Mean



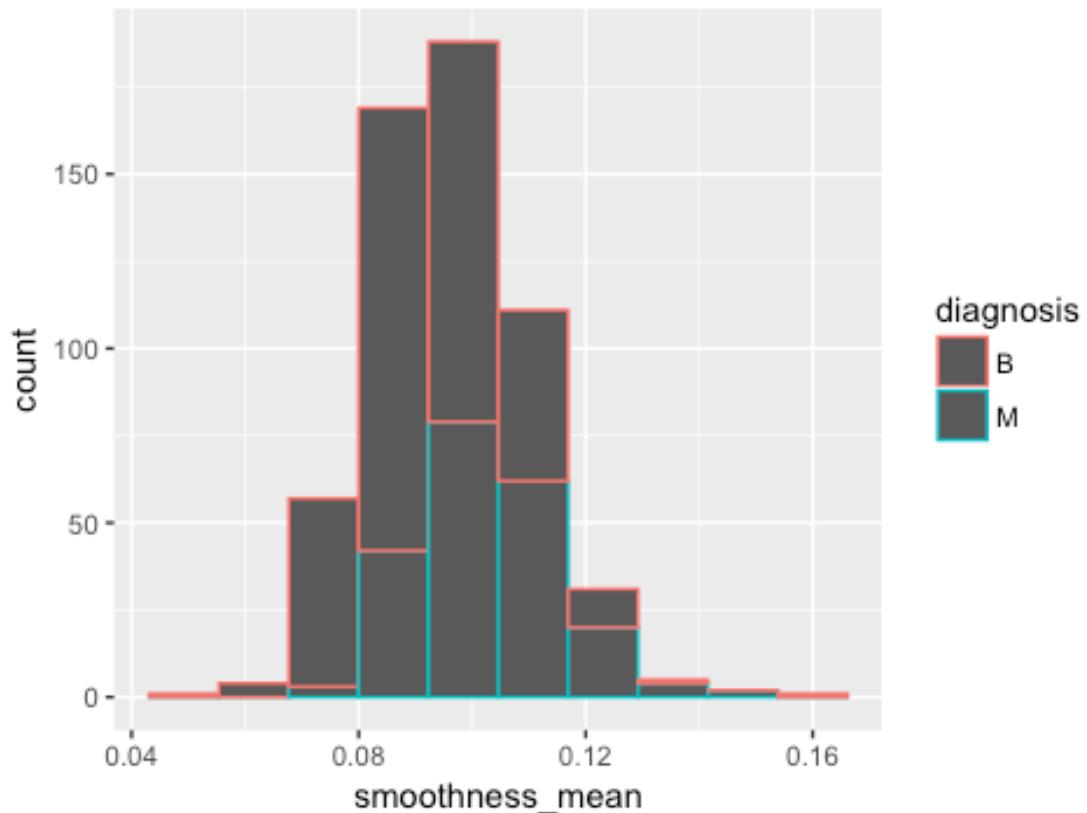
```
#area_mean
hist_area_mean <- ggplot(data = cancer, aes(x = area_mean)) + aes(color=diagnosis) +
  geom_histogram(binwidth = 20) + labs(title = "Histogram Area Mean")
hist_area_mean
```

Histogram Area Mean



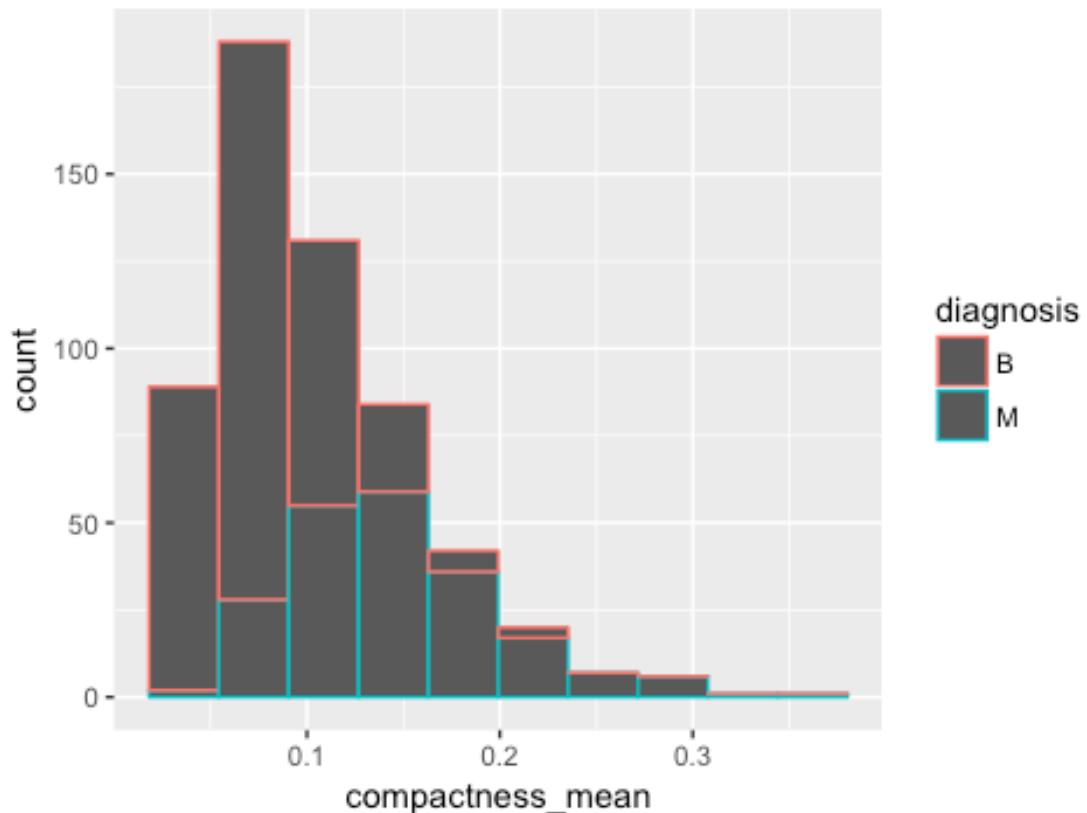
```
#smoothness_mean
hist_smoothness_mean <- ggplot(data = cancer, aes(x = smoothness_mean)) + aes
(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Smoothness Mean")
hist_smoothness_mean
```

Histogram Smoothness Mean



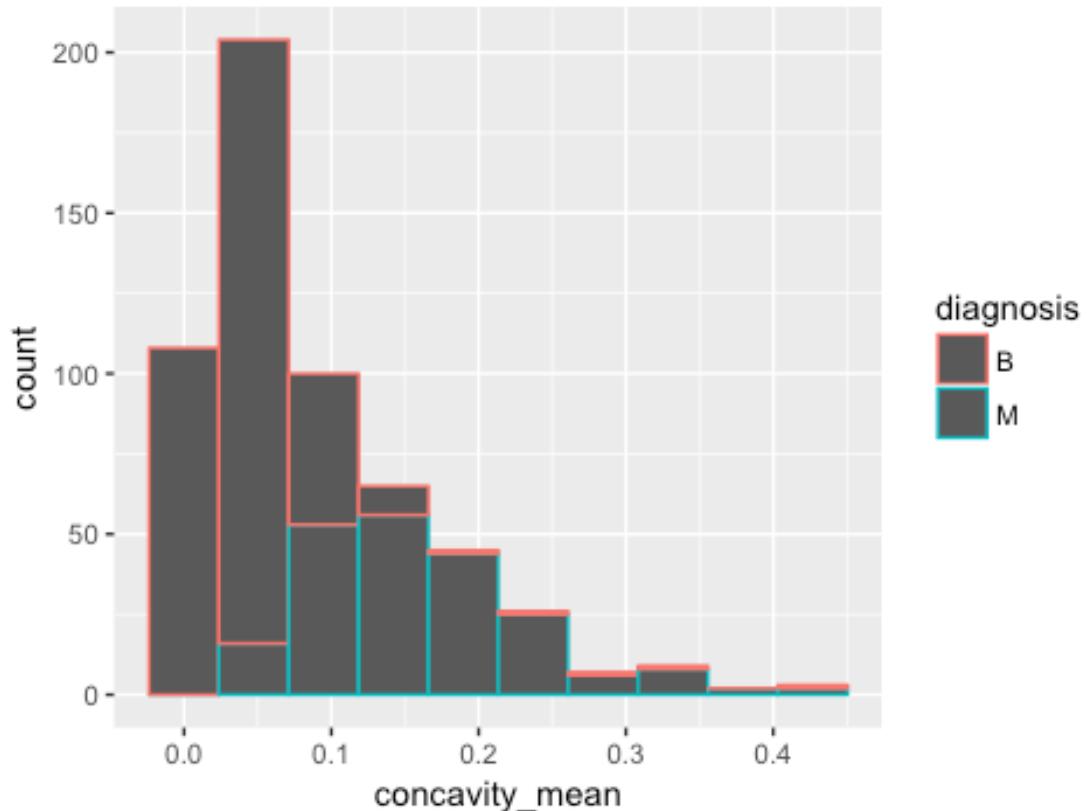
```
#compactness_mean
hist_compactness_mean <- ggplot(data = cancer, aes(x = compactness_mean)) + aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Compactness Mean")
hist_compactness_mean
```

Histogram Compactness Mean



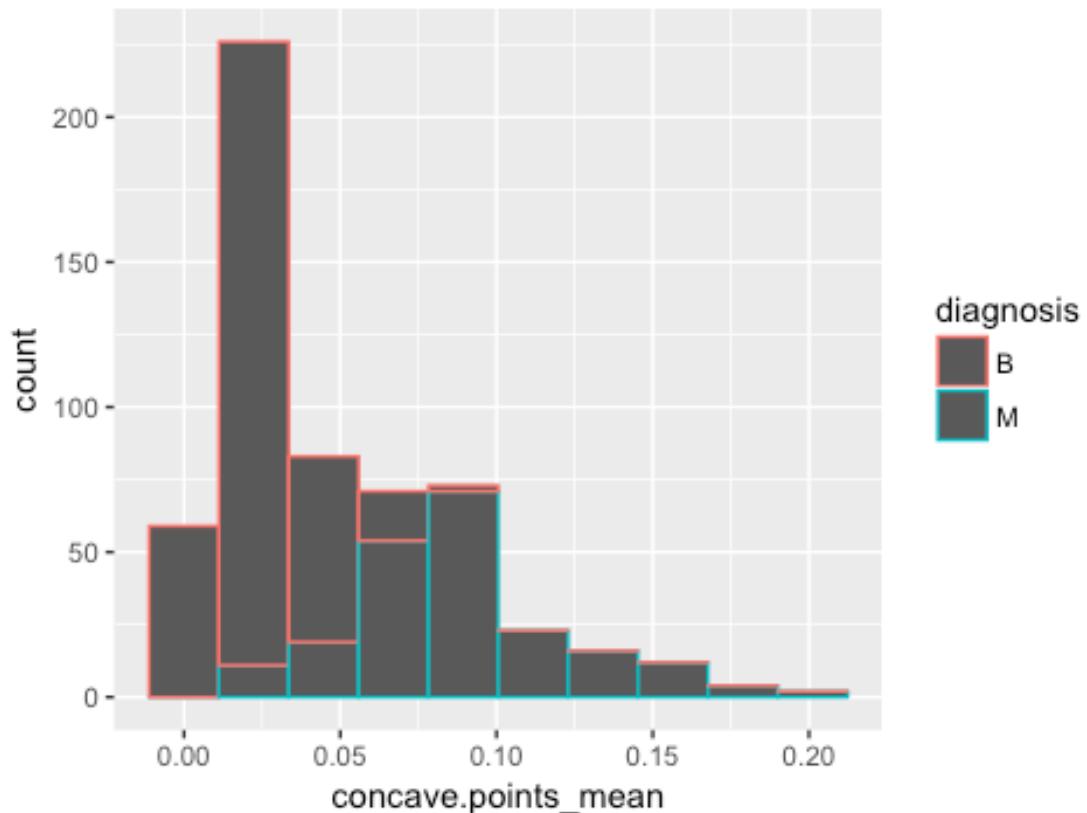
```
#concavity_mean
hist_concavity_mean <- ggplot(data = cancer, aes(x = concavity_mean)) + aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Concavity Mean")
hist_concavity_mean
```

Histogram Concavity Mean



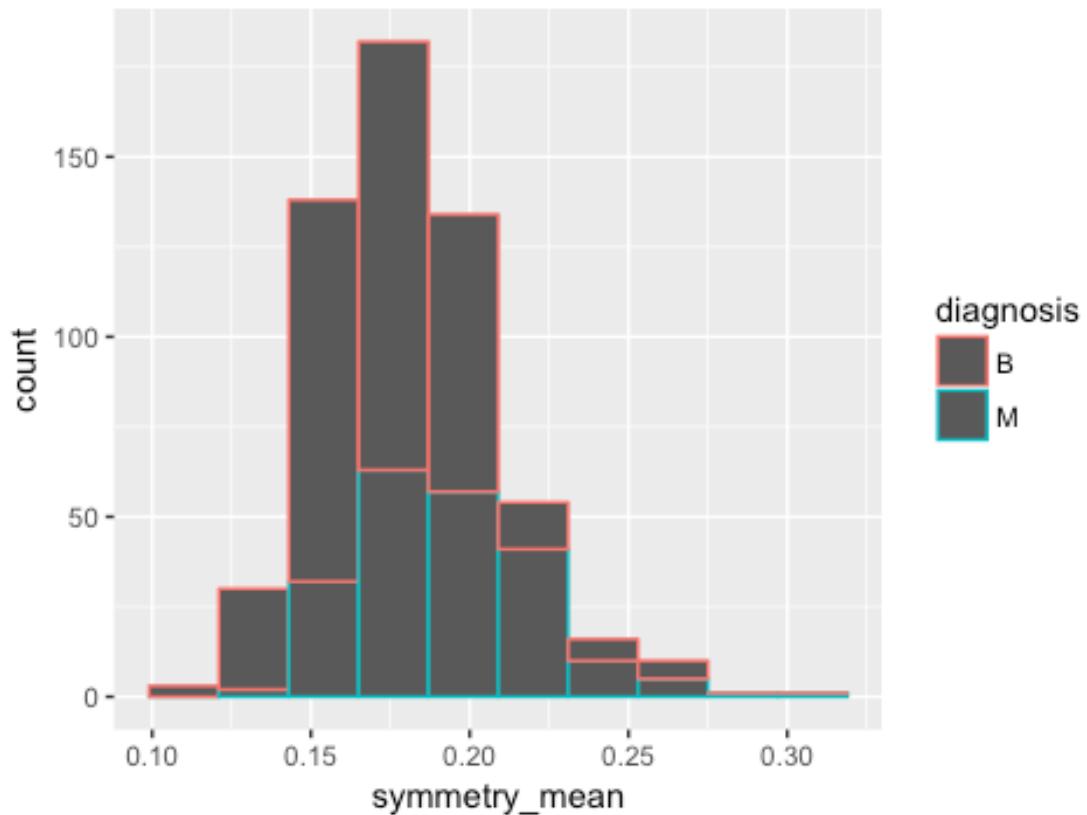
```
#concavepoints_mean
hist_concavepoints_mean <- ggplot(data = cancer, aes(x = concave.points_mean)
) + aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Concave Points Mean")
hist_concavepoints_mean
```

Histogram Concave Points Mean



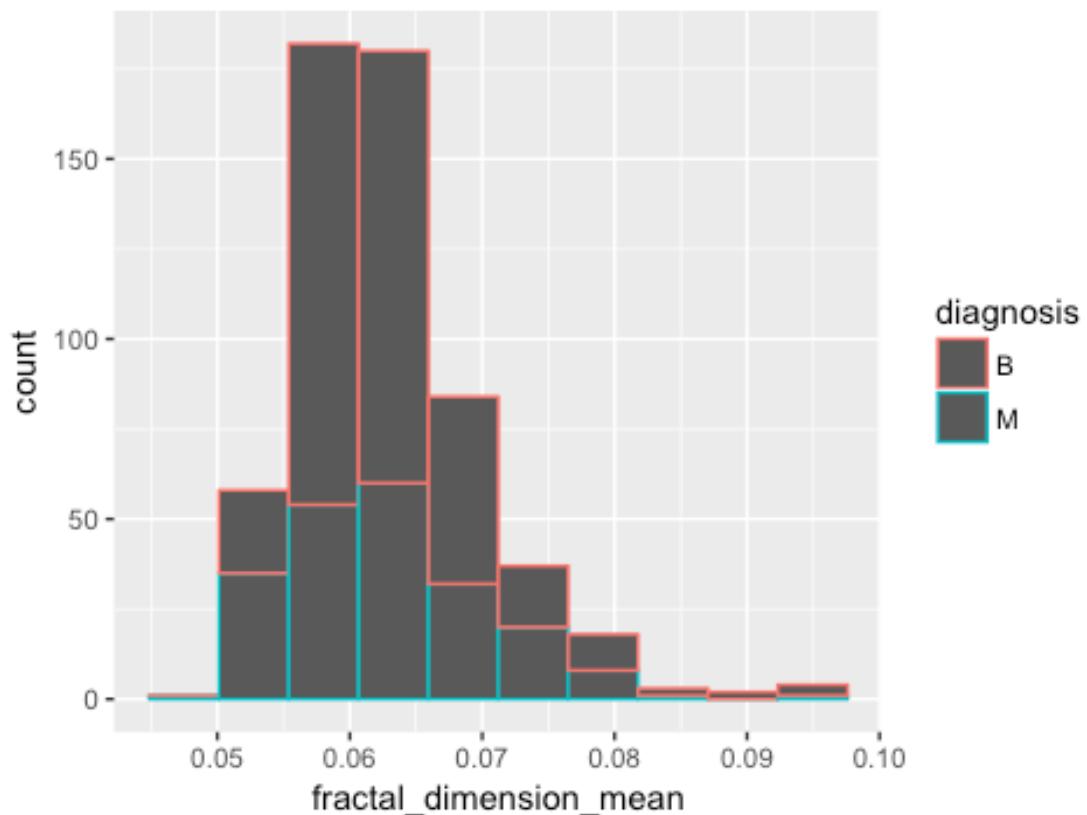
```
#concavepoints_mean
hist_symmetry_mean <- ggplot(data = cancer, aes(x = symmetry_mean)) + aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Symmetry Mean")
hist_symmetry_mean
```

Histogram Symmetry Mean



```
#fractaldim_mean
hist_fractaldim_mean <- ggplot(data = cancer, aes(x = fractal_dimension_mean)
) + aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Fractal Dimension Mean"
)
hist_fractaldim_mean
```

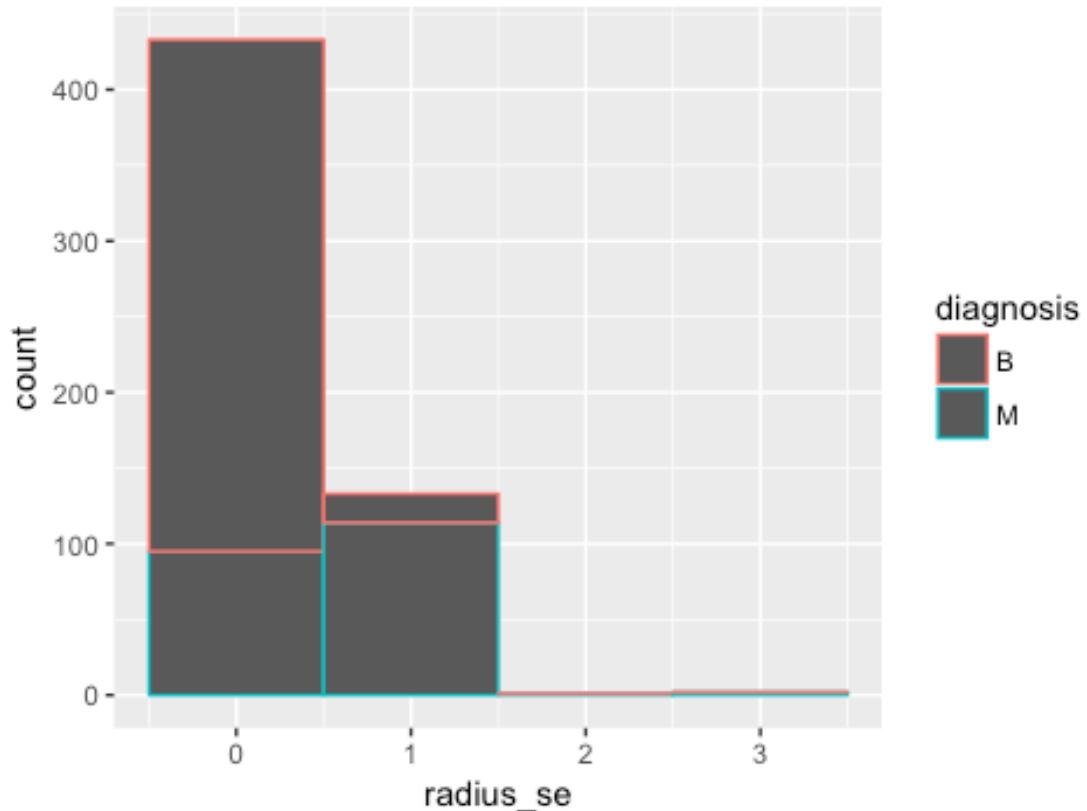
Histogram Fractal Dimension Mean



```
##Histograms of standard error variables
#radius_se
hist_radius_se <- ggplot(data = cancer, aes(x = radius_se)) + aes(color=diagnosis) +
  geom_histogram(binwidth = 1) + labs(title = "Histogram Radius Standard Error")
```

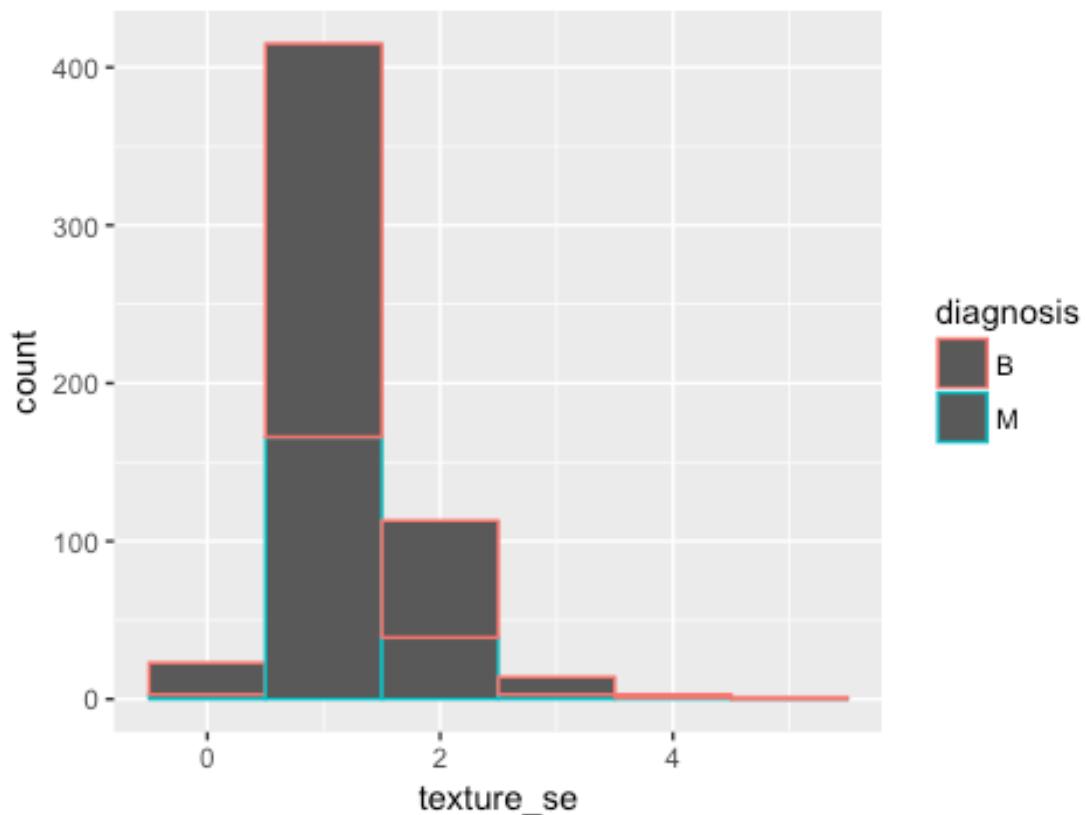
hist_radius_se

Histogram Radius Standard Error



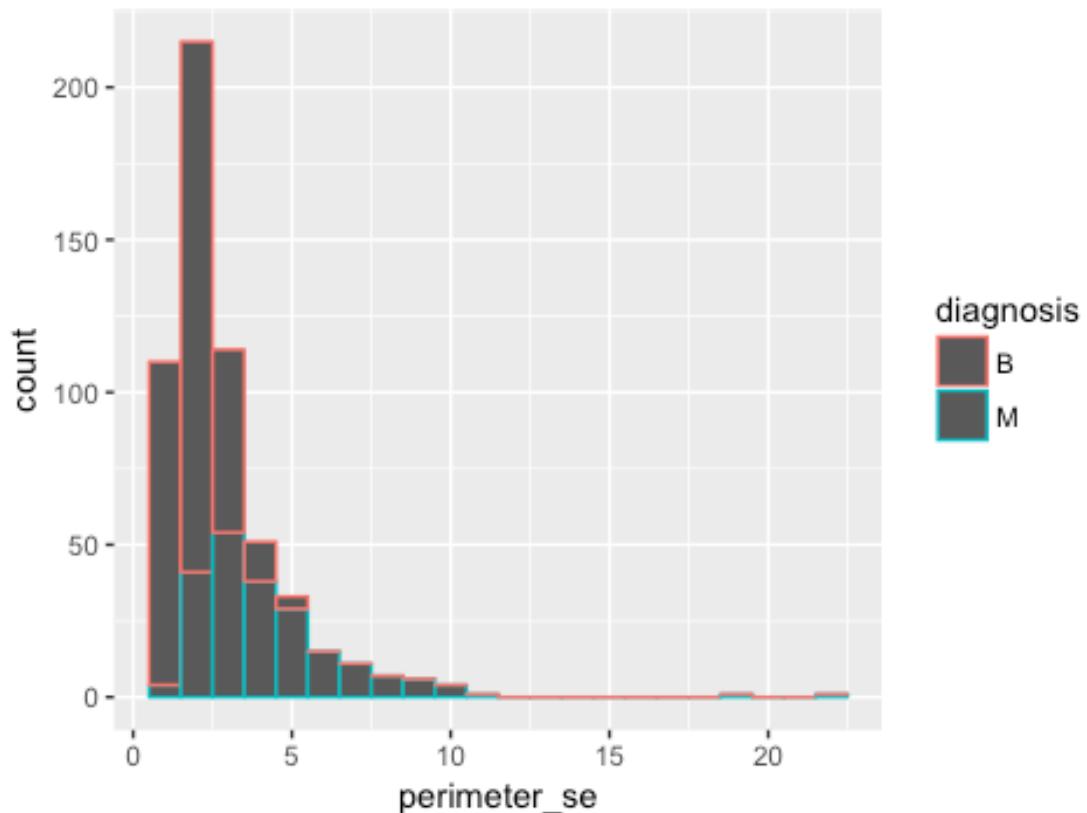
```
#texture_se
hist_texture_se <- ggplot(data = cancer, aes(x = texture_se)) + aes(color=diagnosis) +
  geom_histogram(binwidth = 1) + labs(title = "Histogram Texture Standard Error")
hist_texture_se
```

Histogram Texture Standard Error



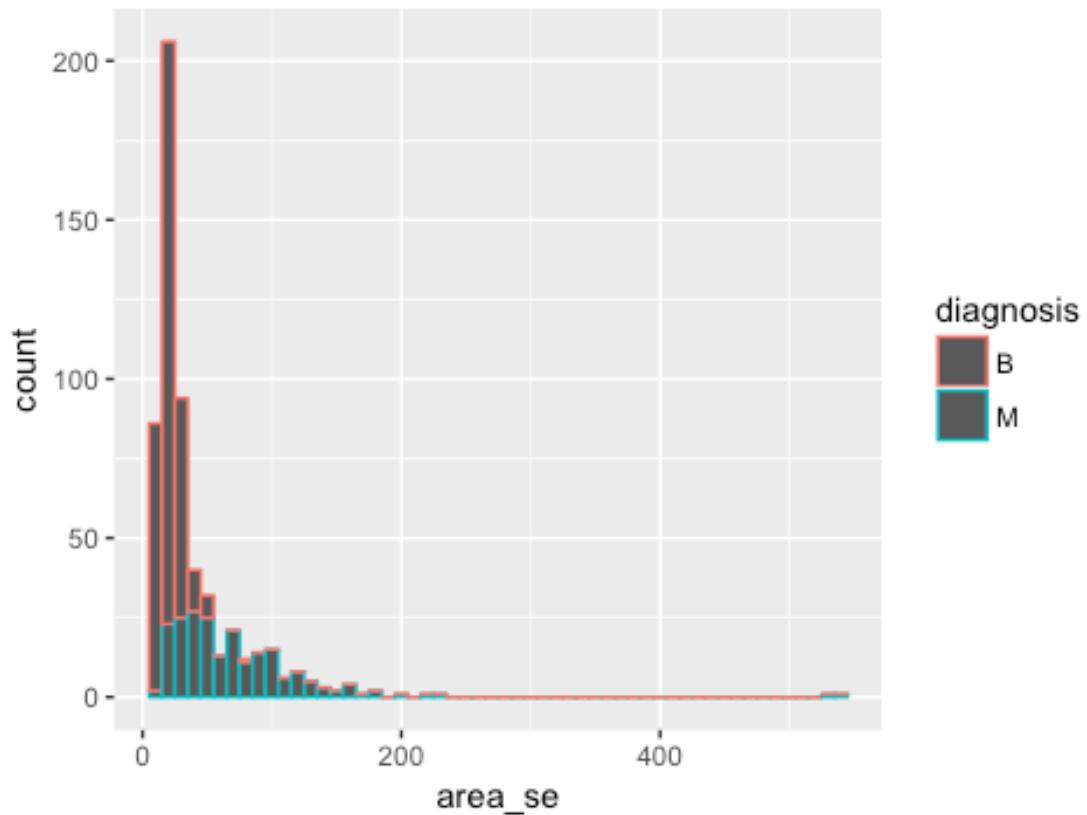
```
#perimeter_se
hist_perimeter_se <- ggplot(data = cancer, aes(x = perimeter_se)) + aes(color = diagnosis) +
  geom_histogram(binwidth = 1) + labs(title = "Histogram Perimeter Standard E
rror")
hist_perimeter_se
```

Histogram Perimeter Standard Error



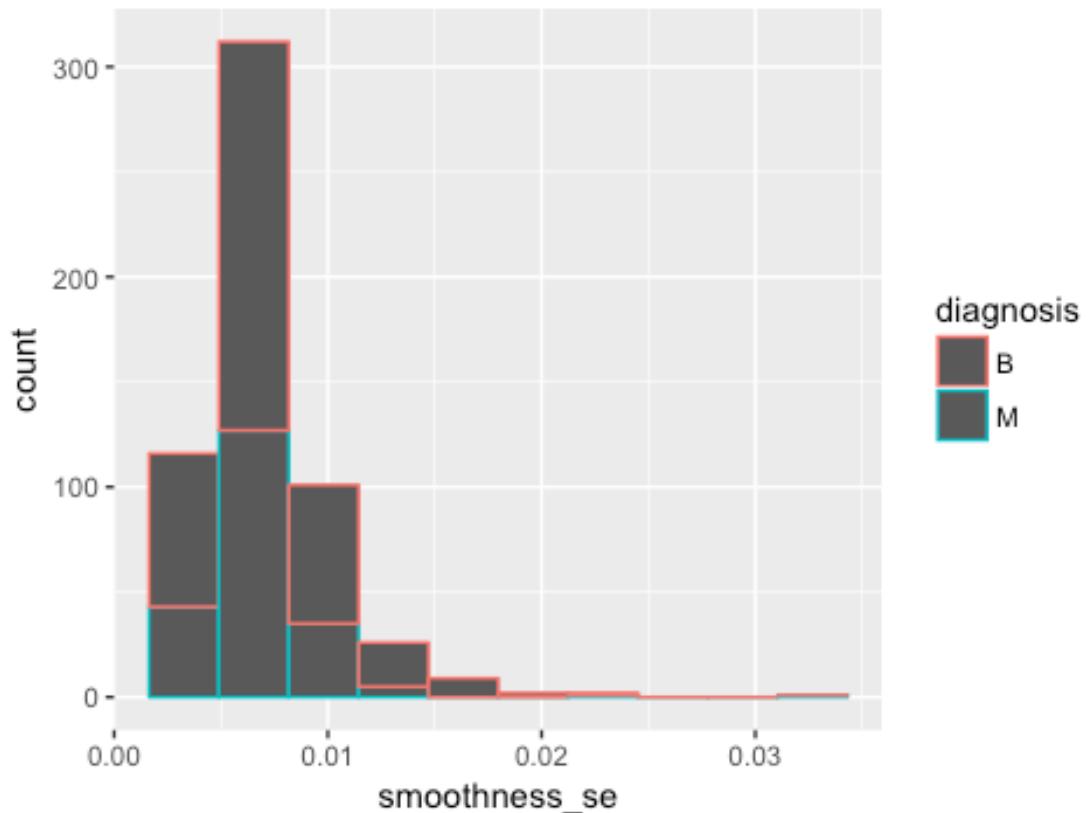
```
#area_se
hist_area_se <- ggplot(data = cancer, aes(x = area_se)) + aes(color=diagnosis)
+ geom_histogram(binwidth = 10) + labs(title = "Histogram Area Standard Error")
hist_area_se
```

Histogram Area Standard Error



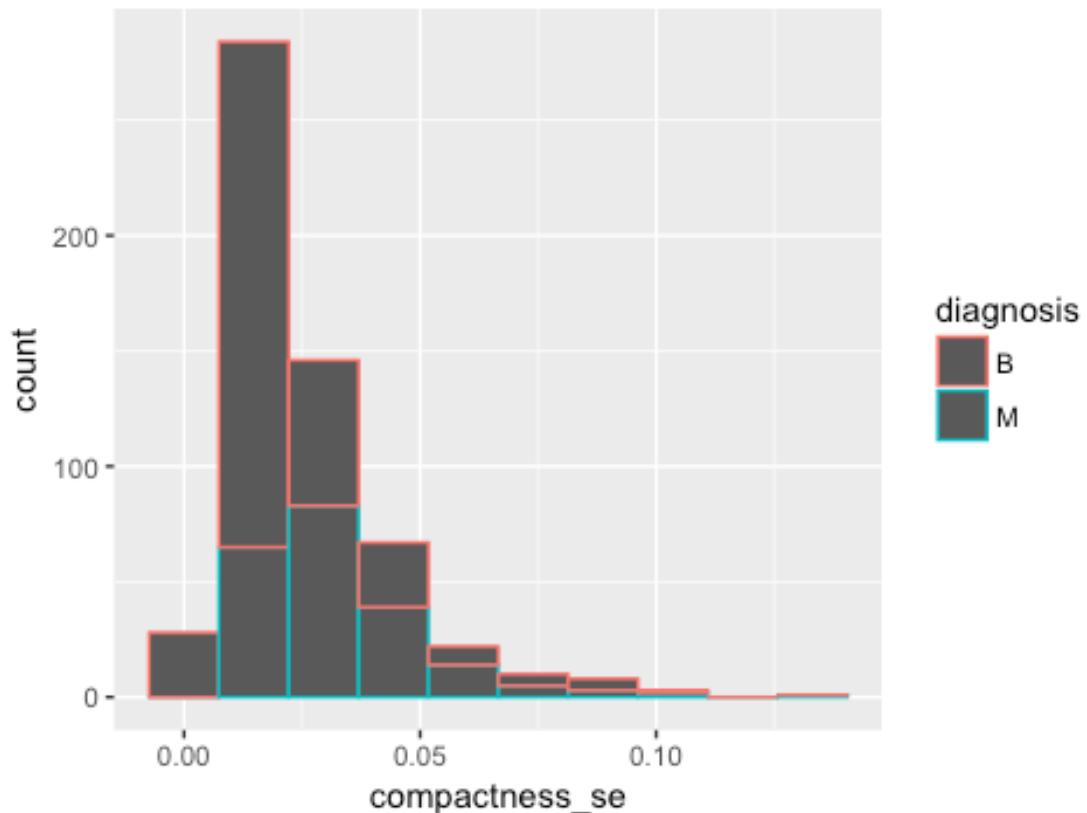
```
#smoothness_se
hist_smoothness_se <- ggplot(data = cancer, aes(x = smoothness_se)) + aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Smoothness Standard Error")
hist_smoothness_se
```

Histogram Smoothness Standard Error



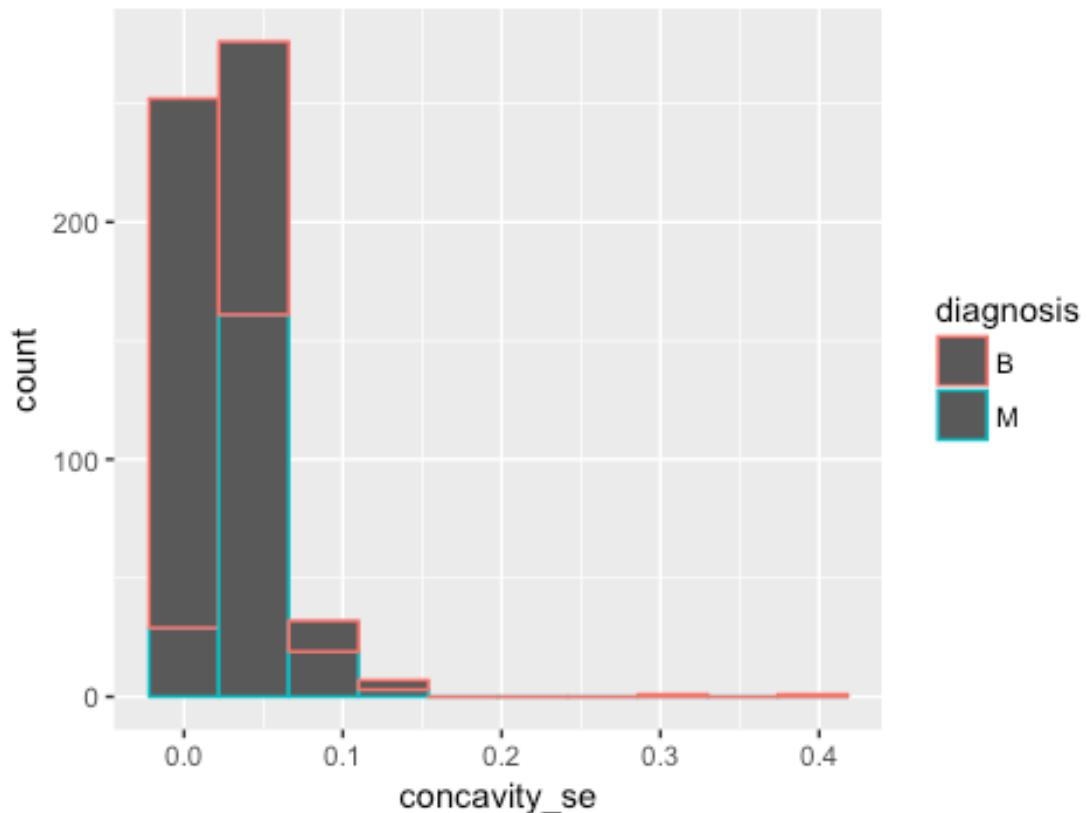
```
#compactness_se
hist_compactness_se <- ggplot(data = cancer, aes(x = compactness_se)) + aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Compactness Standard Error")
hist_compactness_se
```

Histogram Compactness Standard Error

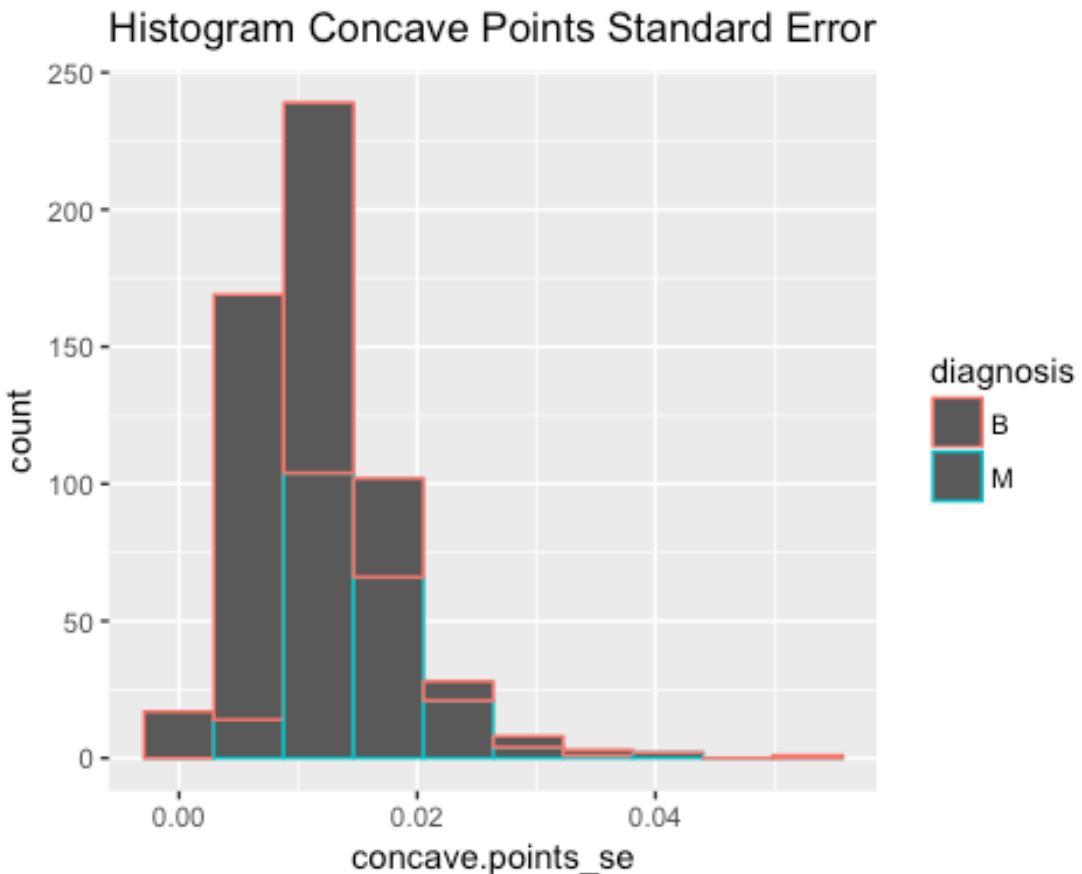


```
#concavity_se
hist_concavity_se <- ggplot(data = cancer, aes(x = concavity_se)) + aes(color = diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Concavity Standard Error")
hist_concavity_se
```

Histogram Concavity Standard Error

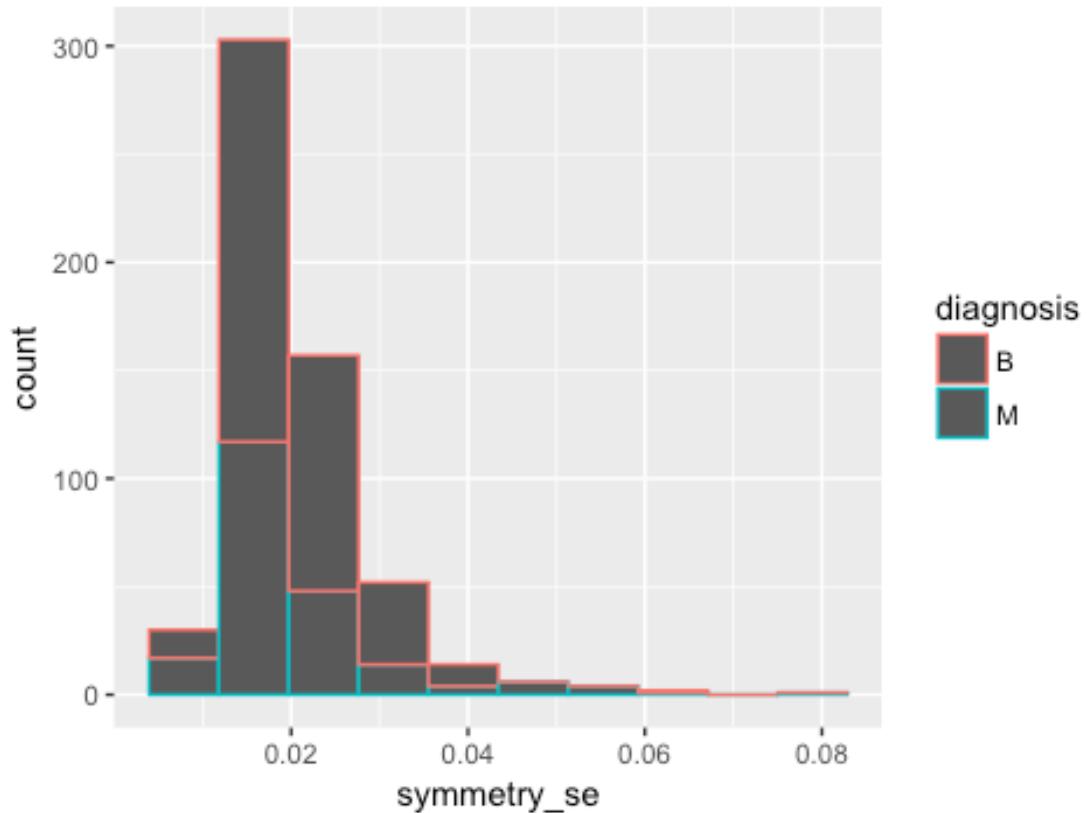


```
#concave_points_se
hist_concavepoints_se <- ggplot(data = cancer, aes(x = concave.points_se)) +
  aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Concave Points Standard
Error")
hist_concavepoints_se
```



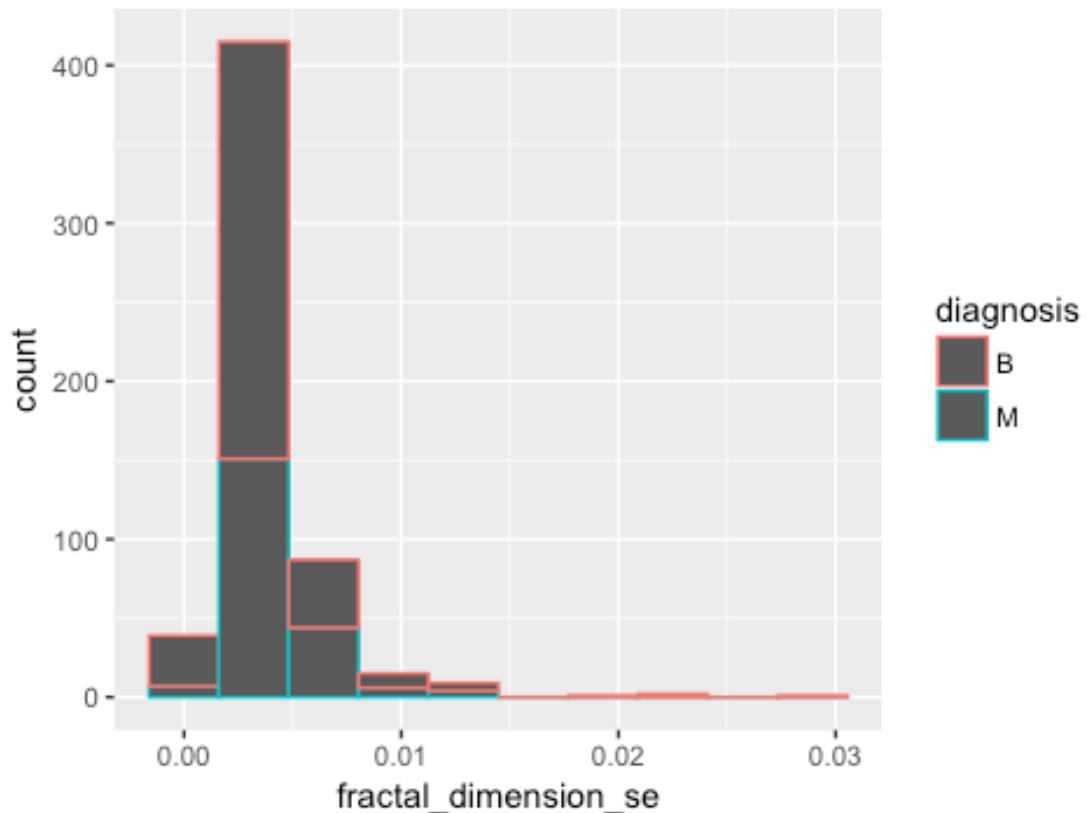
```
#symmetry_se
hist_symmetry_se <- ggplot(data = cancer, aes(x = symmetry_se)) + aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Symmetry Standard Error")
hist_symmetry_se
```

Histogram Symmetry Standard Error



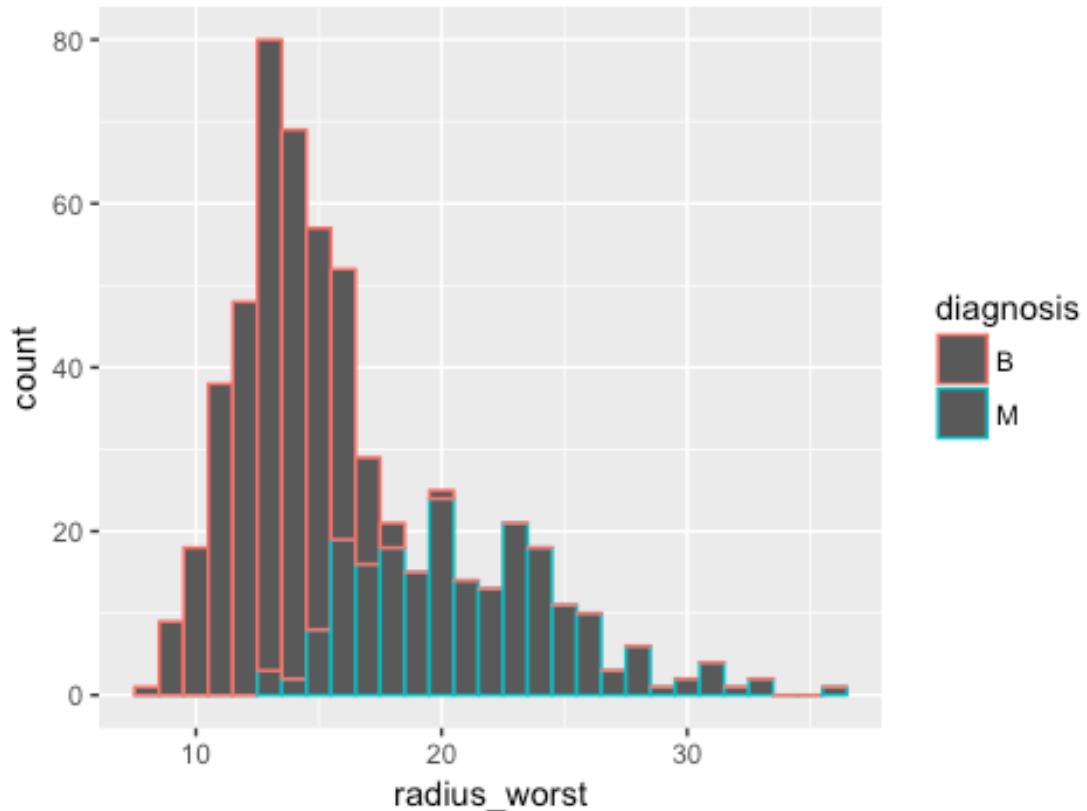
```
#fractaldim_se
hist_fractaldim_se <- ggplot(data = cancer, aes(x = fractal_dimension_se)) +
  aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Fractal Dimension Standard Error")
hist_fractaldim_se
```

Histogram Fractal Dimension Standard Error



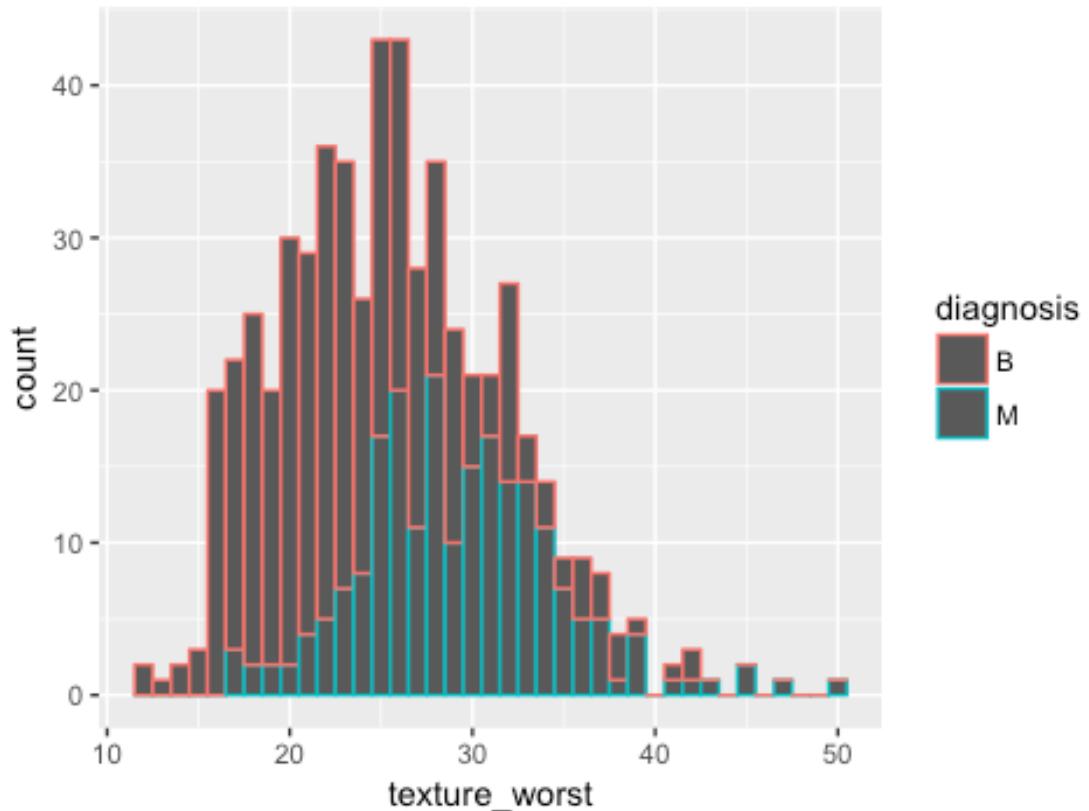
```
##Histograms of worst variables
#radius_worst
hist_radius_worst <- ggplot(data = cancer, aes(x = radius_worst)) + aes(color = diagnosis) +
  geom_histogram(binwidth = 1) + labs(title = "Histogram Radius Maximum")
hist_radius_worst
```

Histogram Radius Maximum



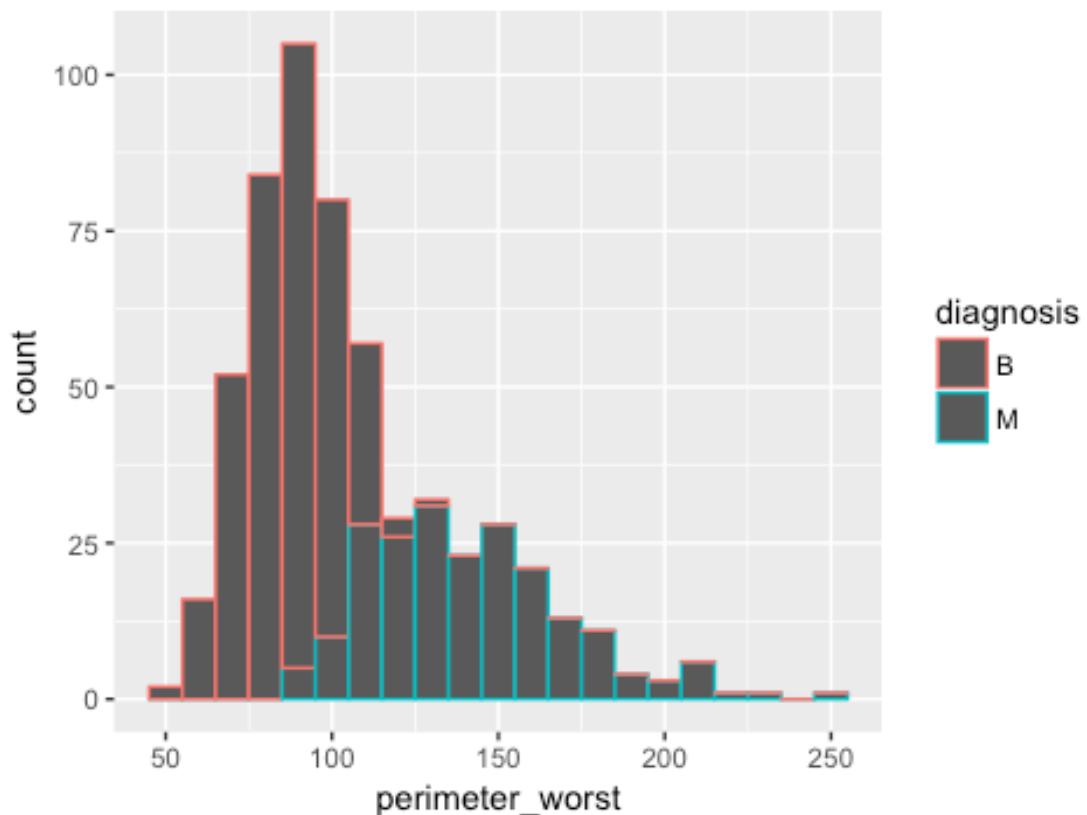
```
#texture_worst
hist_texture_worst <- ggplot(data = cancer, aes(x = texture_worst)) + aes(color=diagnosis) +
  geom_histogram(binwidth = 1) + labs(title = "Histogram Texture Maximum")
hist_texture_worst
```

Histogram Texture Maximum



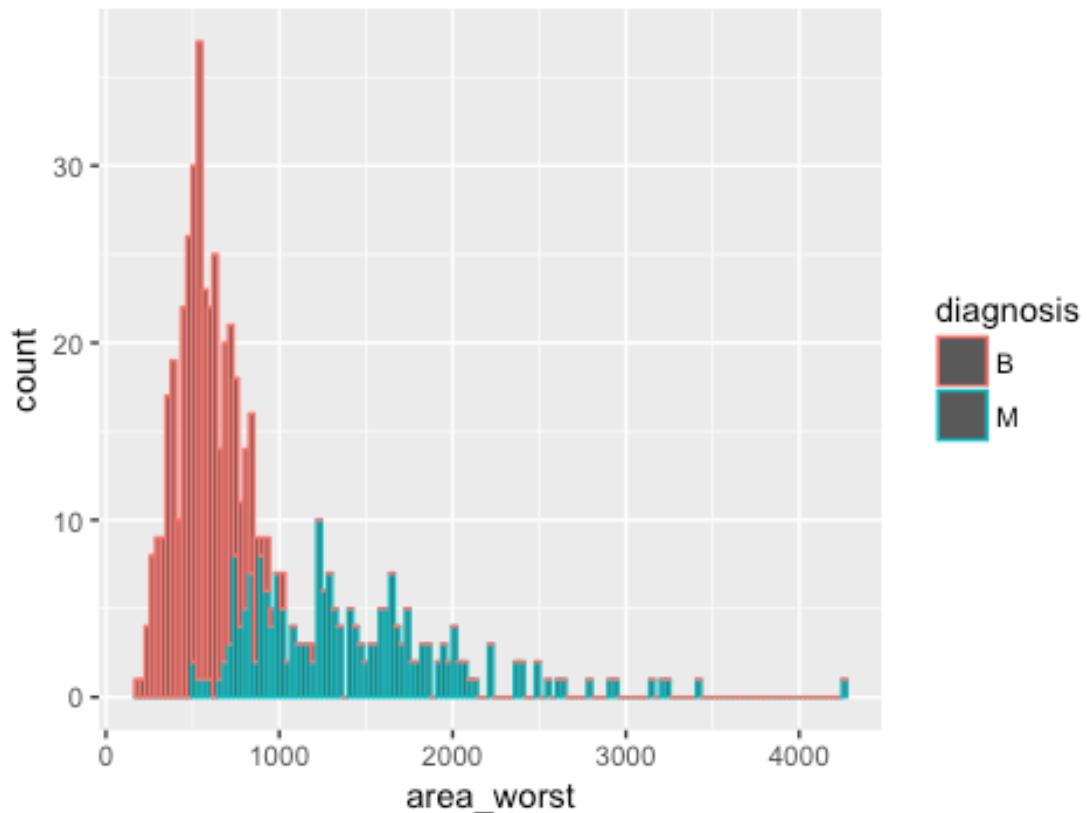
```
#perimeter_worst
hist_perimeter_worst <- ggplot(data = cancer, aes(x = perimeter_worst)) + aes
(color=diagnosis) +
  geom_histogram(binwidth = 10) + labs(title = "Histogram Perimeter Maximum")
hist_perimeter_worst
```

Histogram Perimeter Maximum



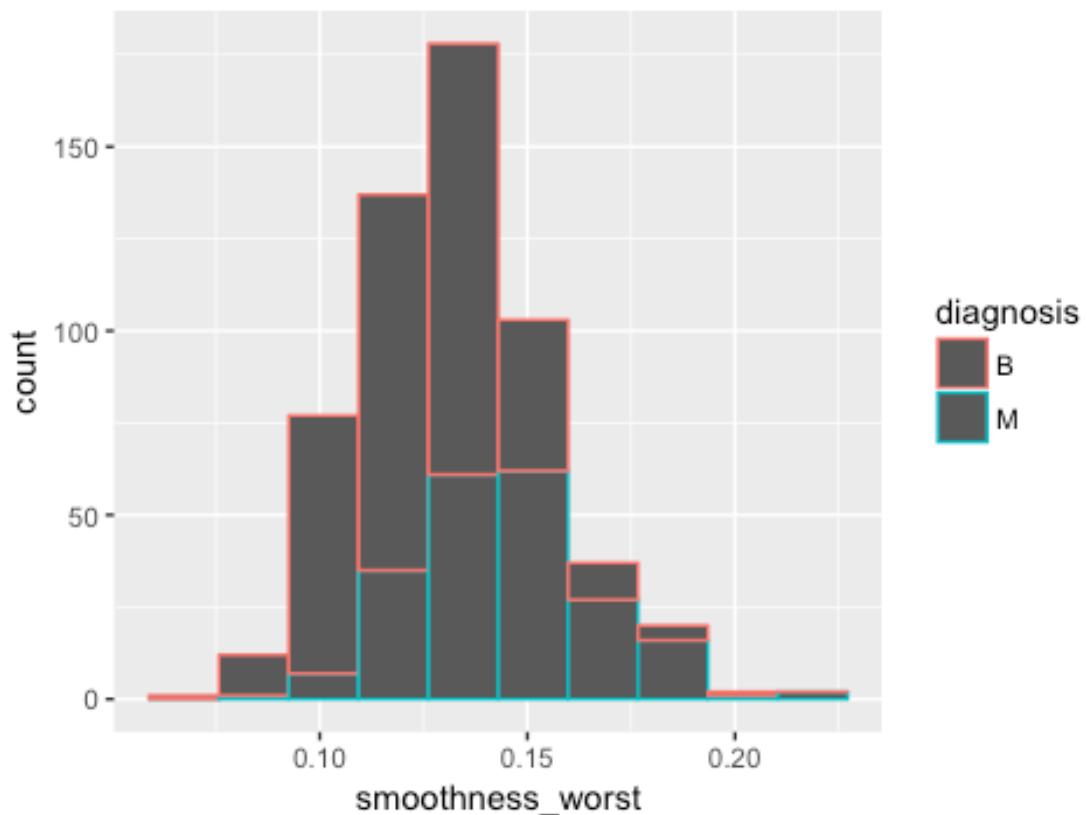
```
#area_worst
hist_area_worst <- ggplot(data = cancer, aes(x = area_worst)) + aes(color=diagnosis) +
  geom_histogram(binwidth = 30) + labs(title = "Histogram Area Maximum")
hist_area_worst
```

Histogram Area Maximum



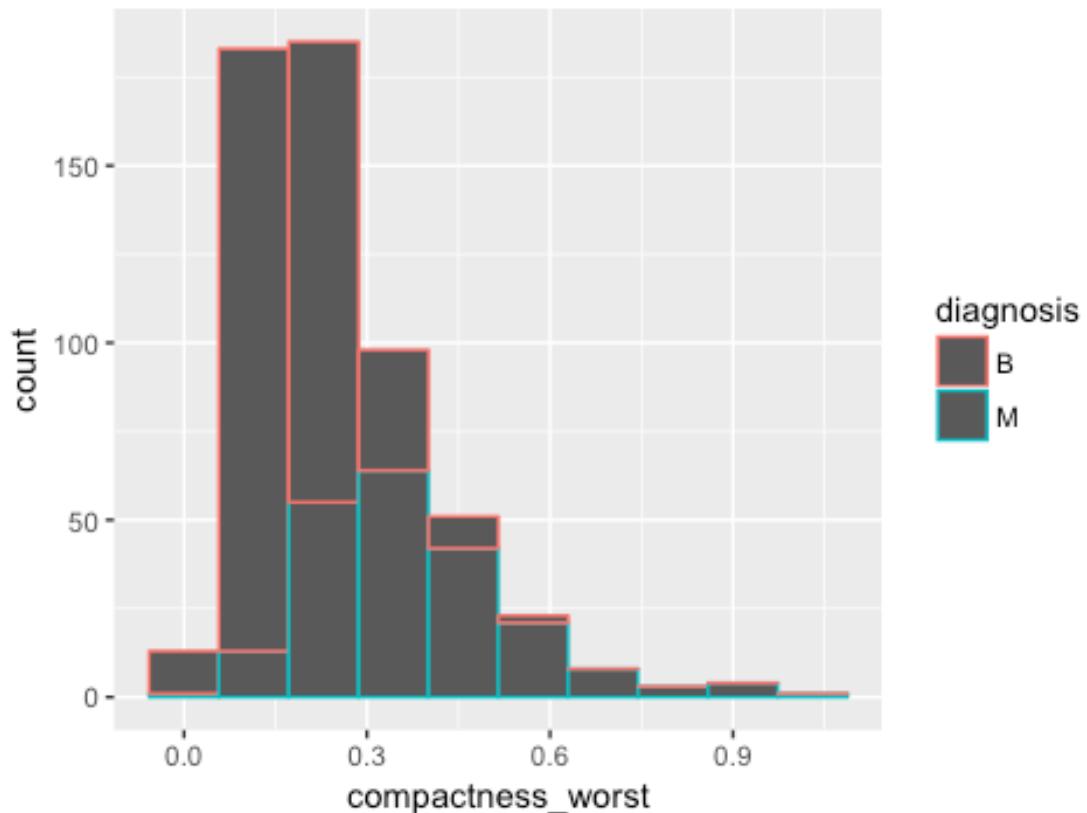
```
#smoothness_worst
hist_smoothness_worst <- ggplot(data = cancer, aes(x = smoothness_worst)) + aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Smoothness Maximum")
hist_smoothness_worst
```

Histogram Smoothness Maximum



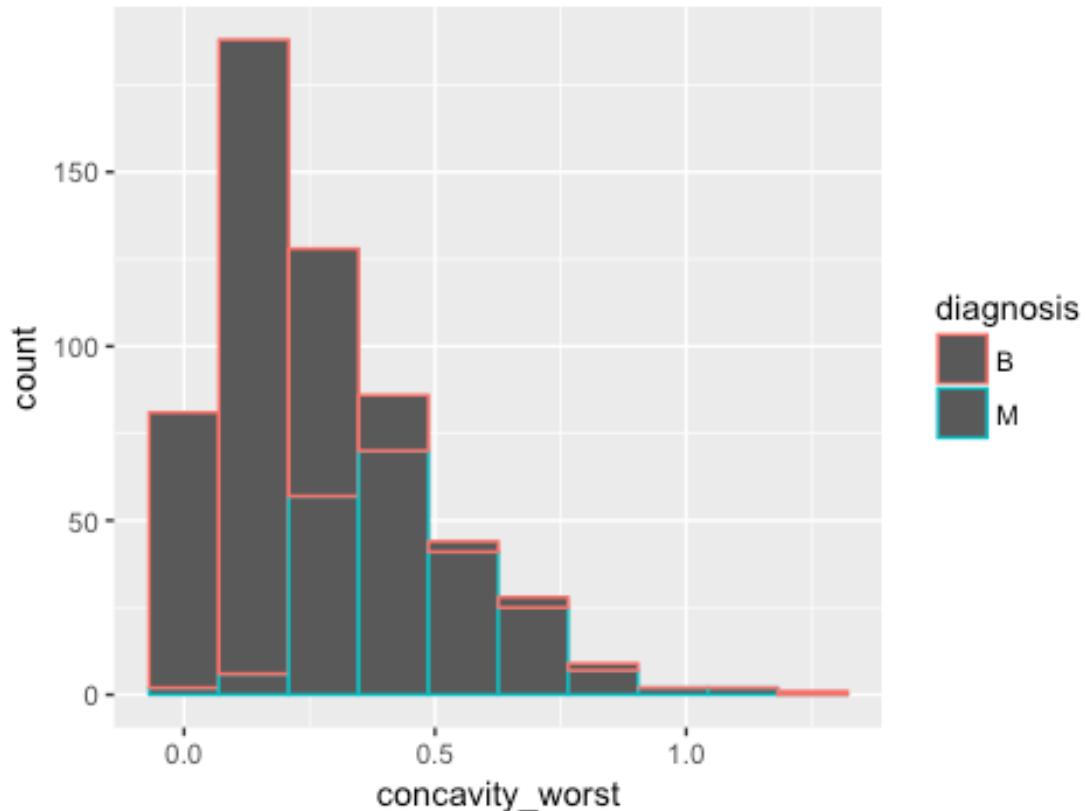
```
#compactness_worst
hist_compactness_worst <- ggplot(data = cancer, aes(x = compactness_worst)) +
  aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Compactness Maximum")
hist_compactness_worst
```

Histogram Compactness Maximum

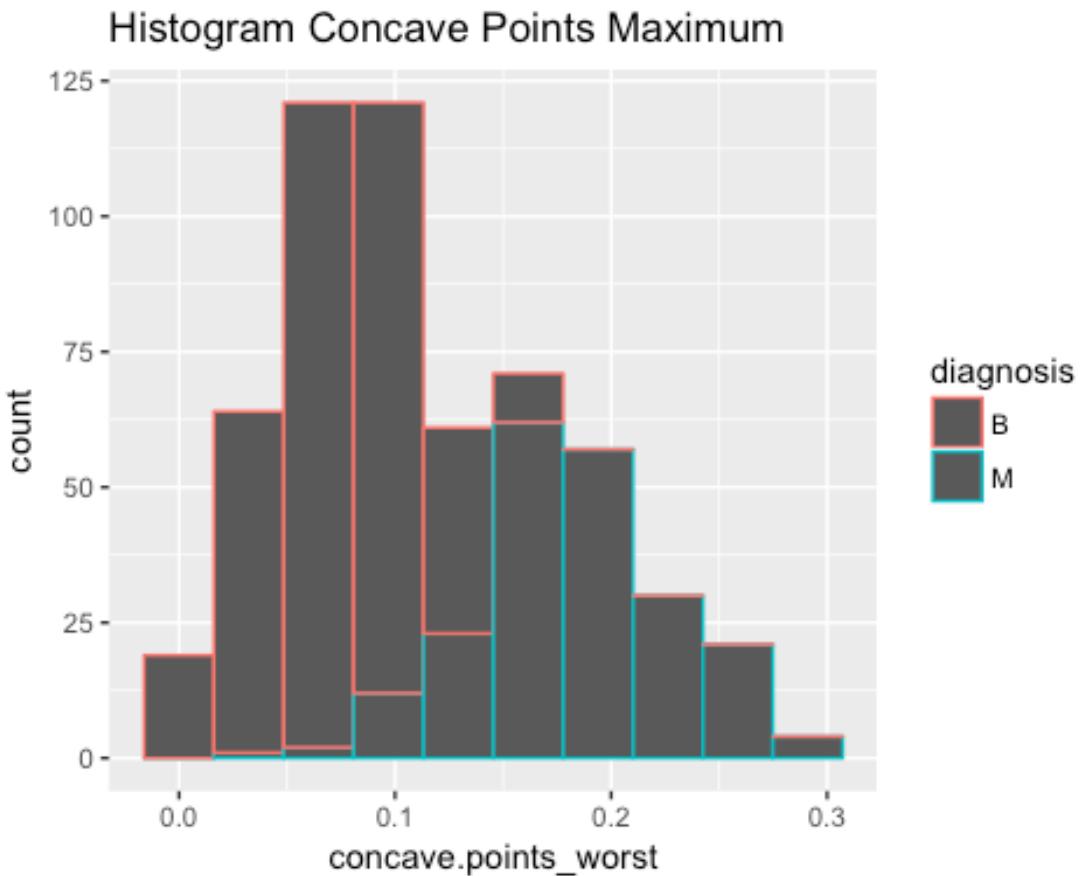


```
#concavity_worst
hist_concavity_worst <- ggplot(data = cancer, aes(x = concavity_worst)) + aes
(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Concavity Maximum")
hist_concavity_worst
```

Histogram Concavity Maximum

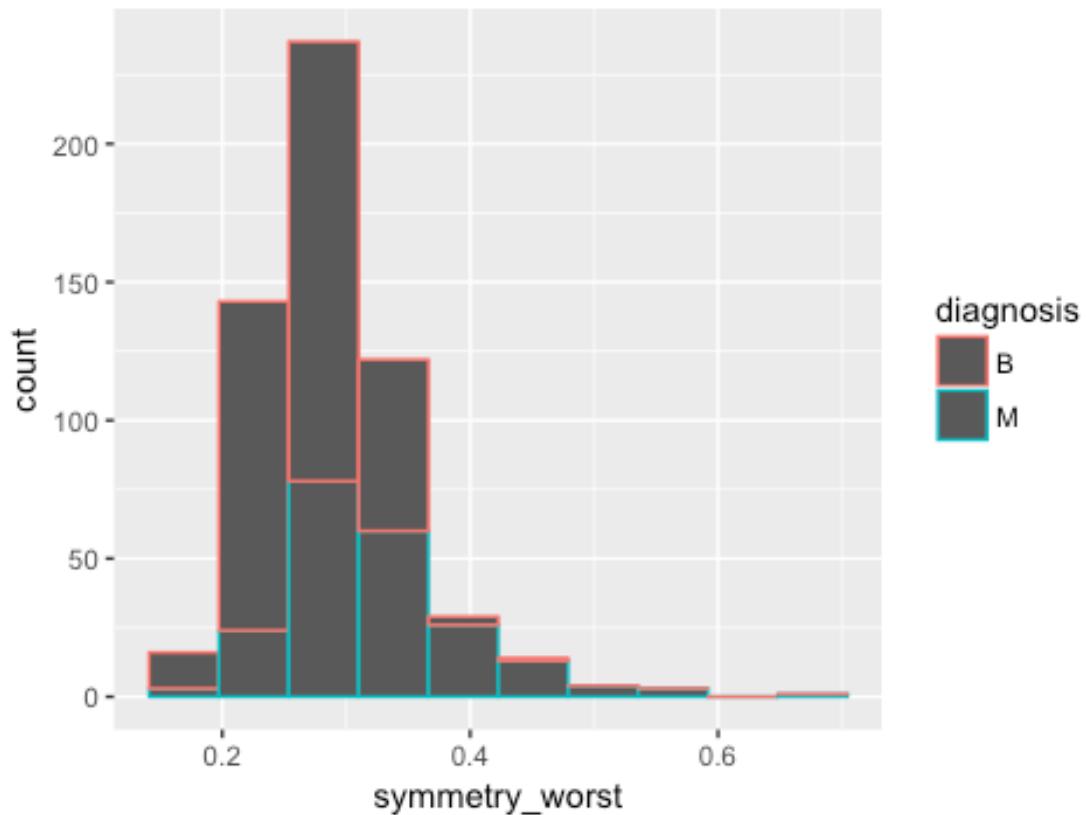


```
#concavepoints_worst
hist_concavepoints_worst <- ggplot(data = cancer, aes(x = concave.points_worst)) + aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Concave Points Maximum")
hist_concavepoints_worst
```



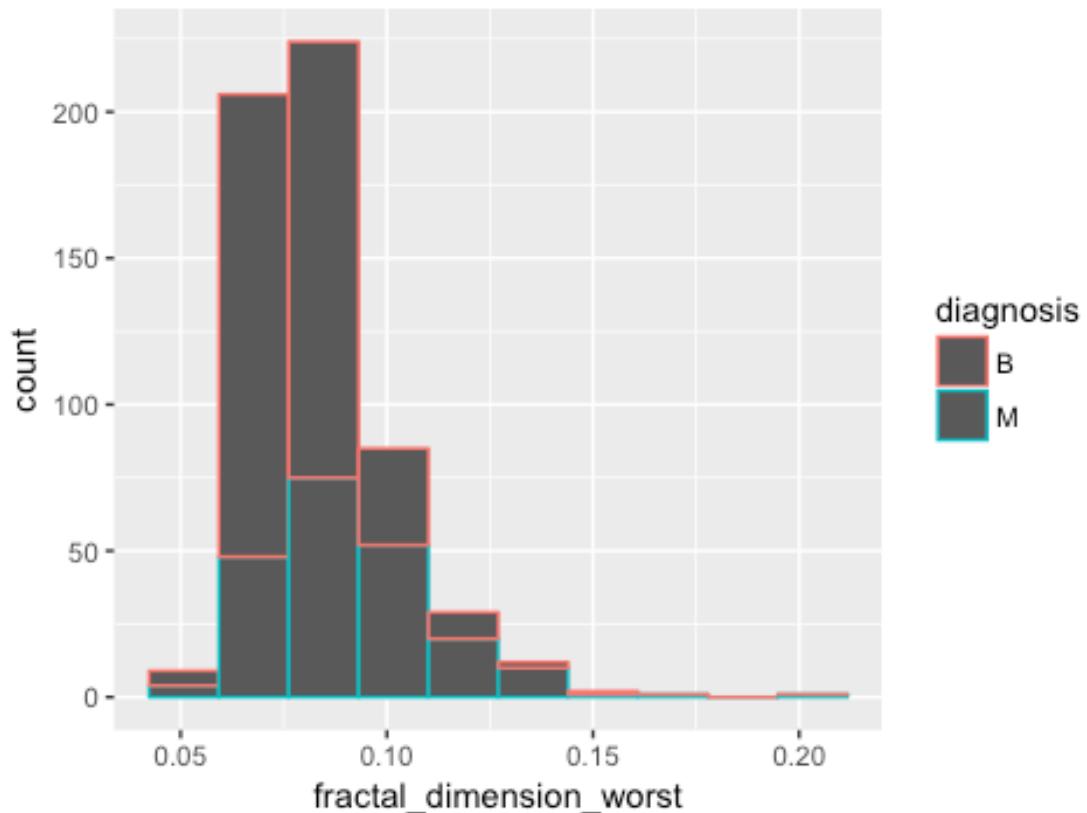
```
#symmetry_worst
hist_symmetry_worst <- ggplot(data = cancer, aes(x = symmetry_worst)) + aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Symmetry Maximum")
hist_symmetry_worst
```

Histogram Symmetry Maximum



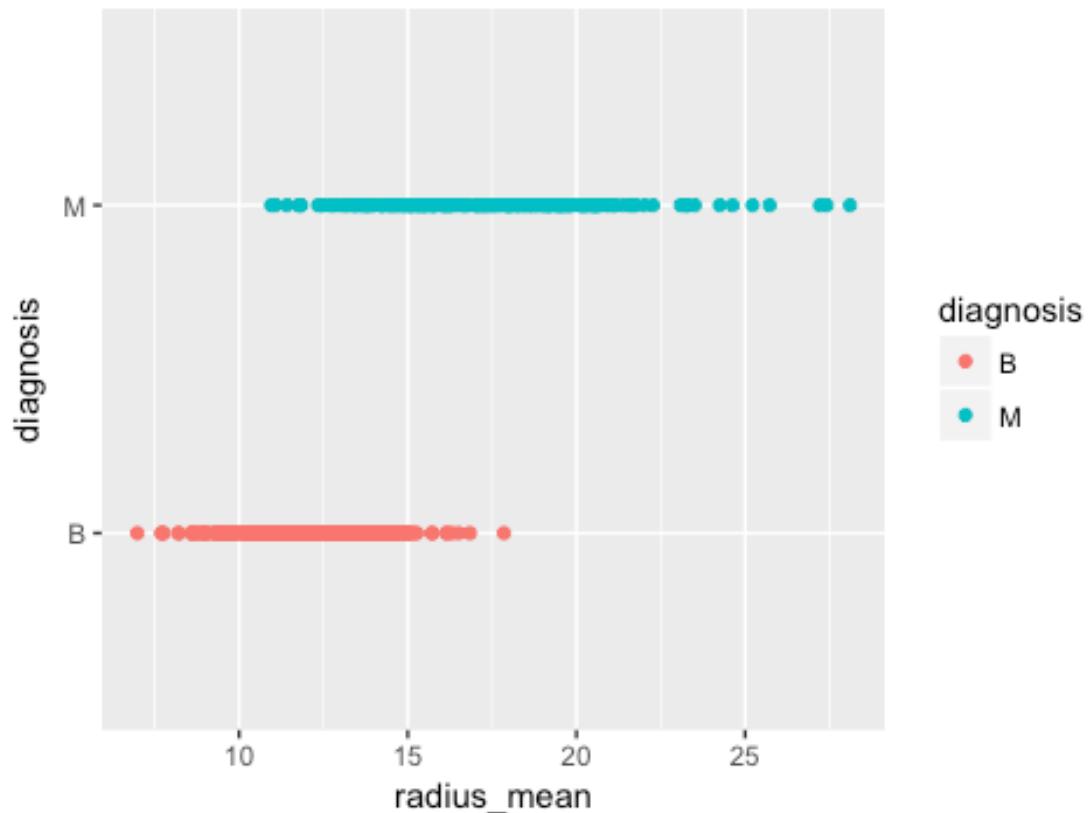
```
#fractaldim_worst
hist_fractaldim_worst <- ggplot(data = cancer, aes(x = fractal_dimension_worst)) + aes(color=diagnosis) +
  geom_histogram(bins = 10) + labs(title = "Histogram Fractal Dimension Maximum")
hist_fractaldim_worst
```

Histogram Fractal Dimension Maximum



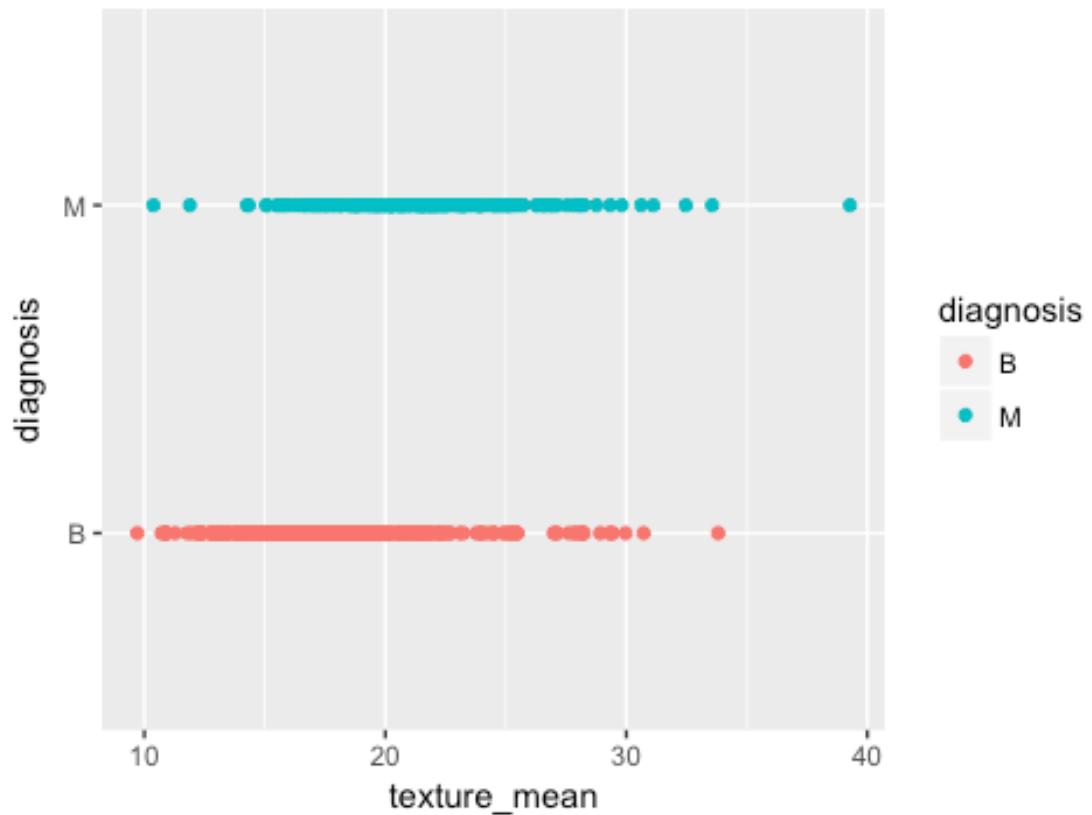
```
#SCATTERPLOTS
##Scatterplots of mean variables
#radius_mean
sc_radius_mean <- ggplot(data = cancer, aes(x = radius_mean, y = diagnosis))
+ aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Radius Mean")
sc_radius_mean
```

Scatterplot Radius Mean



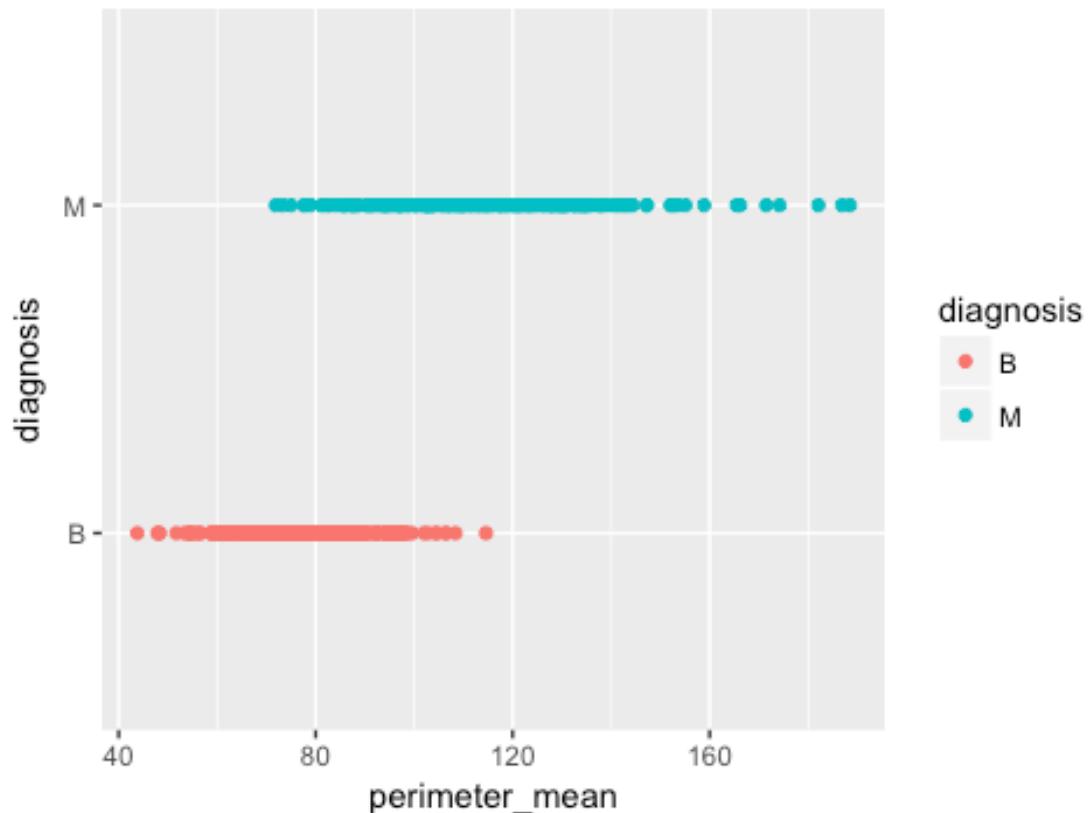
```
#texture_mean
sc_texture_mean <- ggplot(data = cancer, aes(x = texture_mean, y = diagnosis) +
  aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Texture Mean")
sc_texture_mean
```

Scatterplot Texture Mean



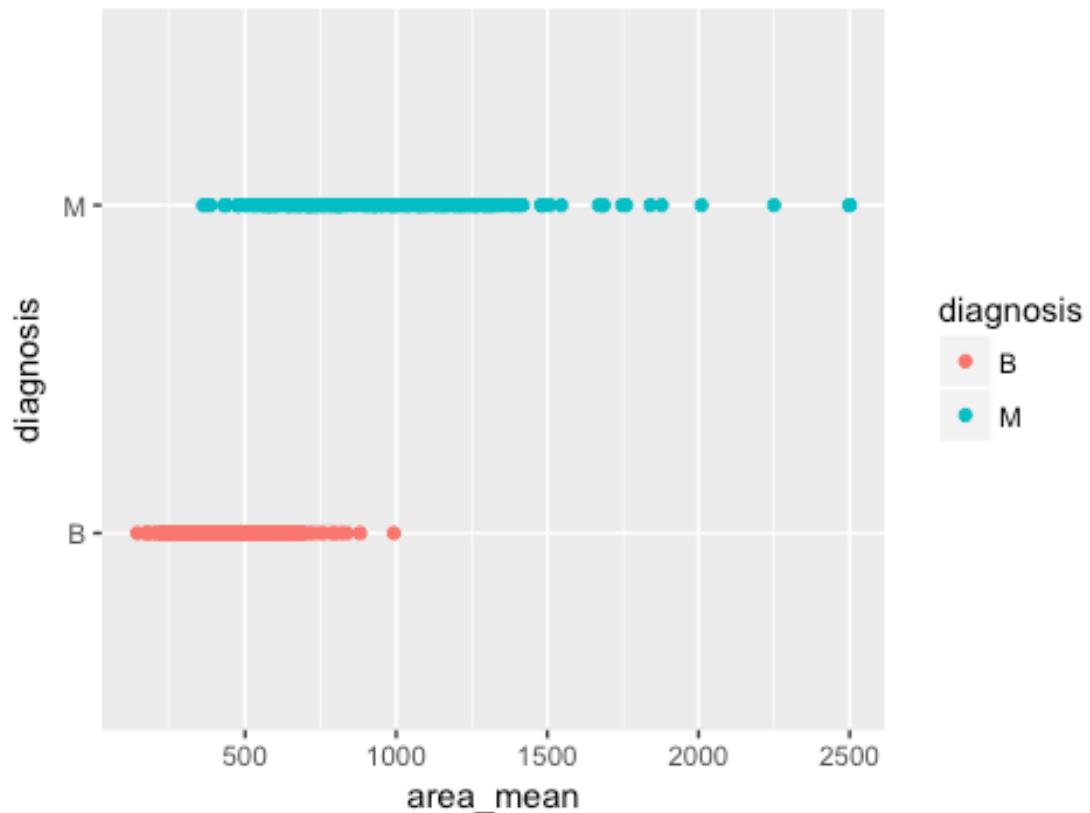
```
#perimeter_mean
sc_perimeter_mean <- ggplot(data = cancer, aes(x = perimeter_mean, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Perimeter Mean")
sc_perimeter_mean
```

Scatterplot Perimeter Mean



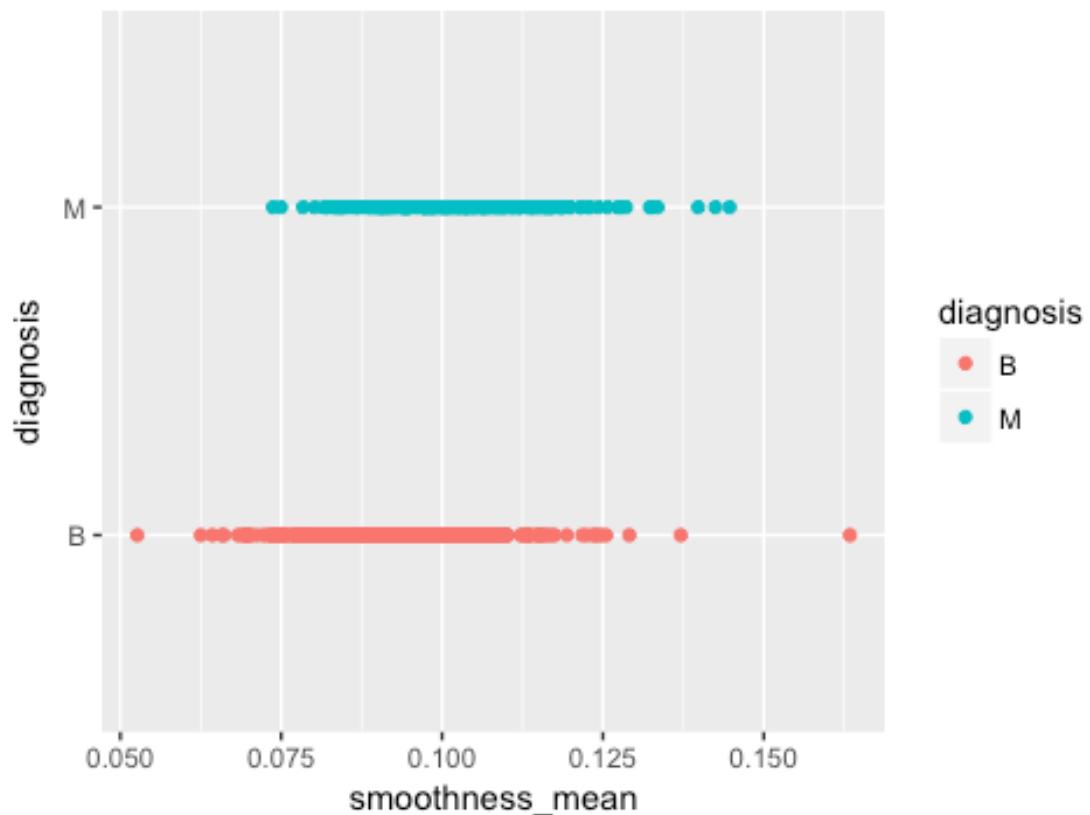
```
#area_mean
sc_area_mean <- ggplot(data = cancer, aes(x = area_mean, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Area Mean")
sc_area_mean
```

Scatterplot Area Mean



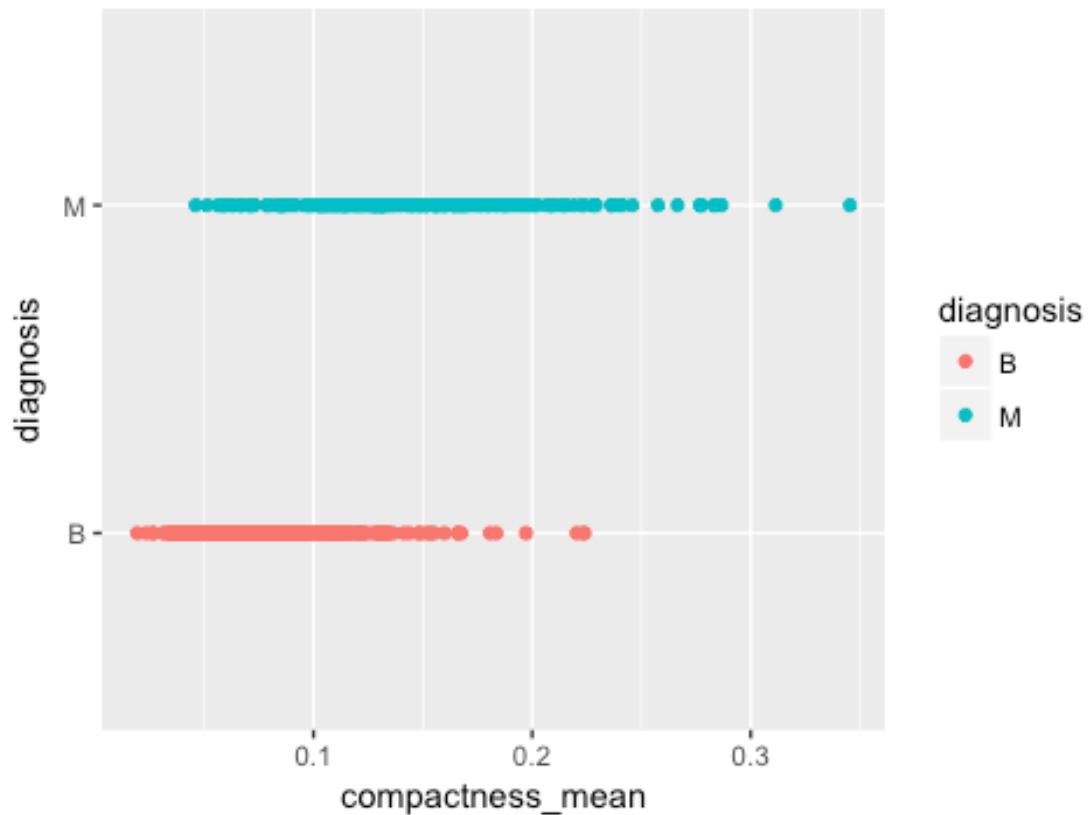
```
#smoothness_mean
sc_smoothness_mean <- ggplot(data = cancer, aes(x = smoothness_mean, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Smoothness Mean")
sc_smoothness_mean
```

Scatterplot Smoothness Mean



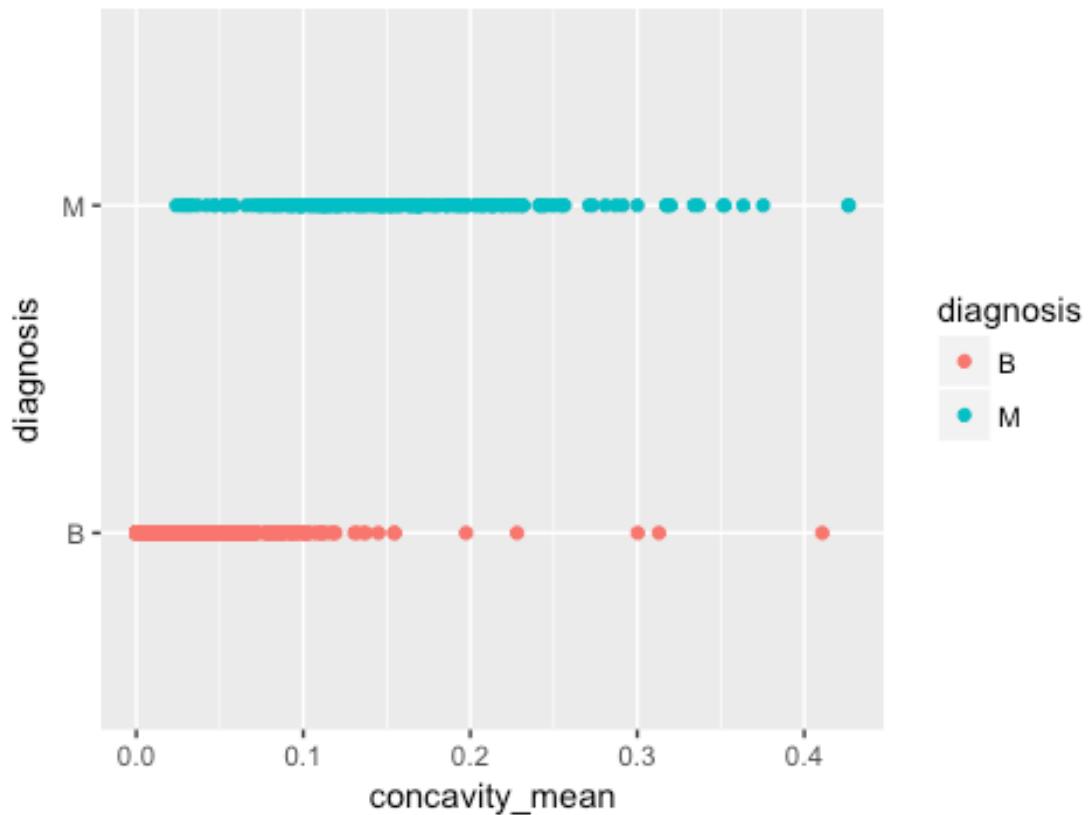
```
#compactness_mean
sc_compactness_mean <- ggplot(data = cancer, aes(x = compactness_mean, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Compactness Mean")
sc_compactness_mean
```

Scatterplot Compactness Mean



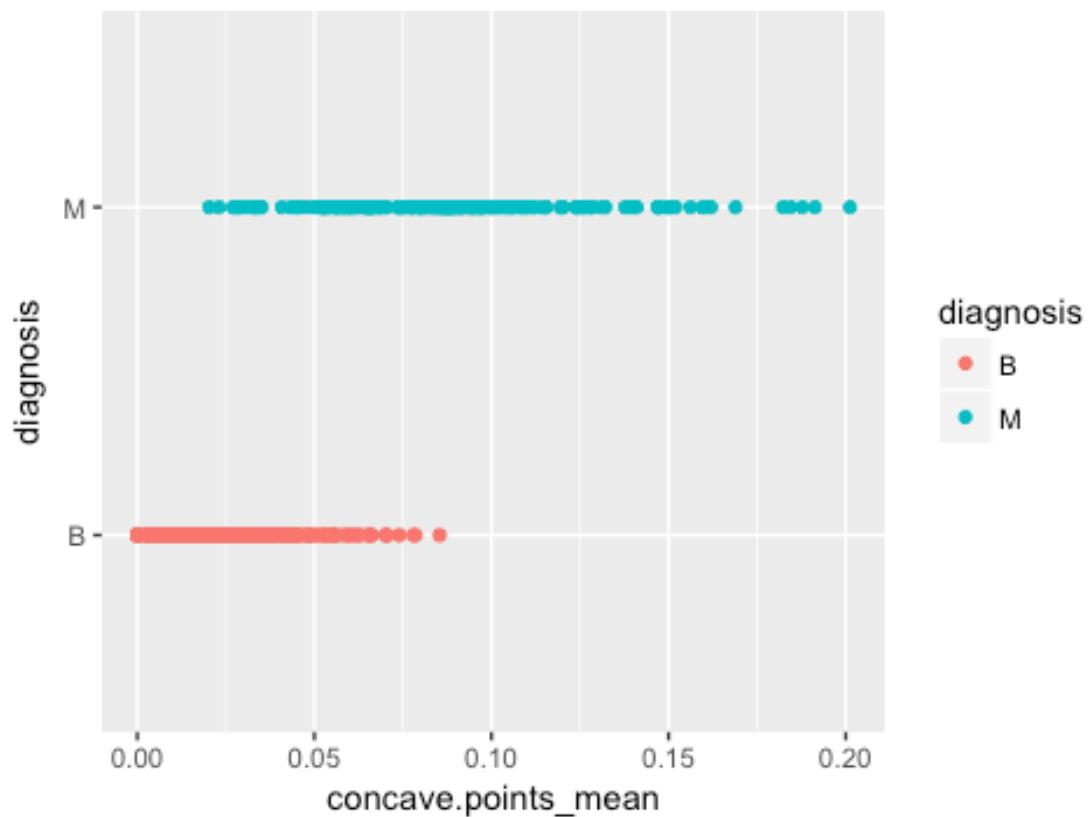
```
#concavity_mean
sc_concavity_mean <- ggplot(data = cancer, aes(x = concavity_mean, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Concavity Mean")
sc_concavity_mean
```

Scatterplot Concavity Mean



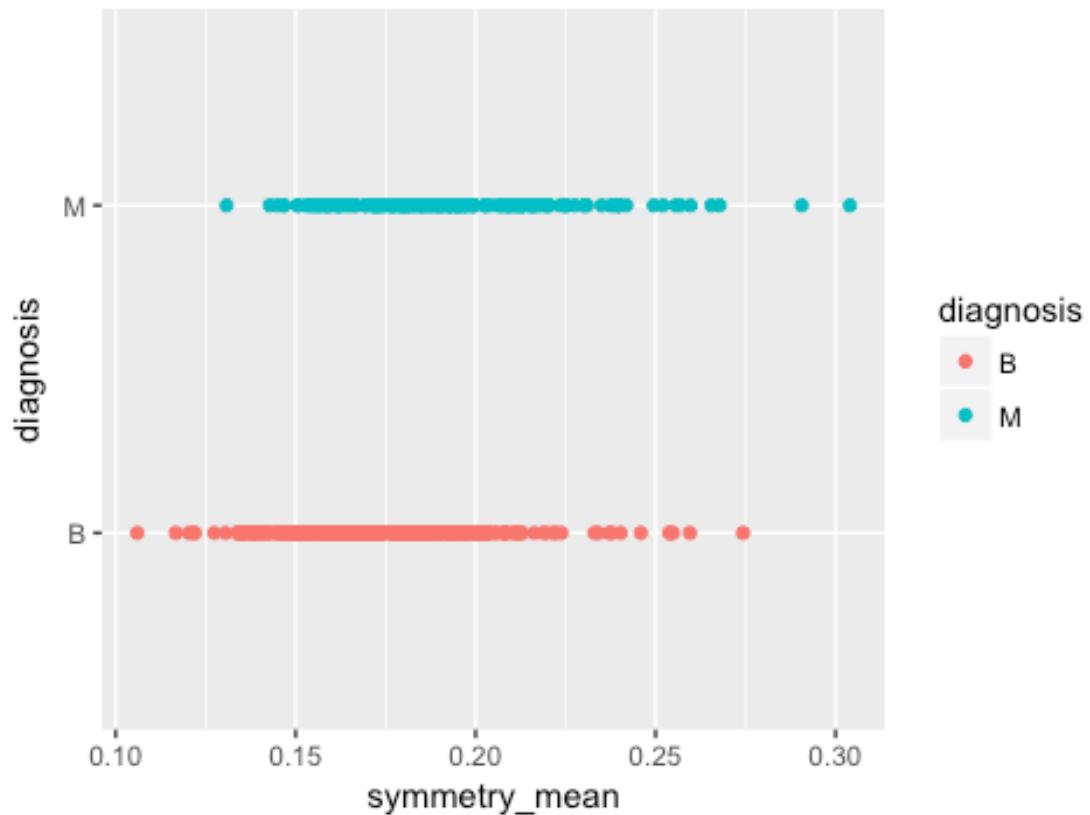
```
#concavepoints_mean
sc_concavepoints_mean <- ggplot(data = cancer, aes(x = concave.points_mean, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Concave Points Mean")
sc_concavepoints_mean
```

Scatterplot Concave Points Mean



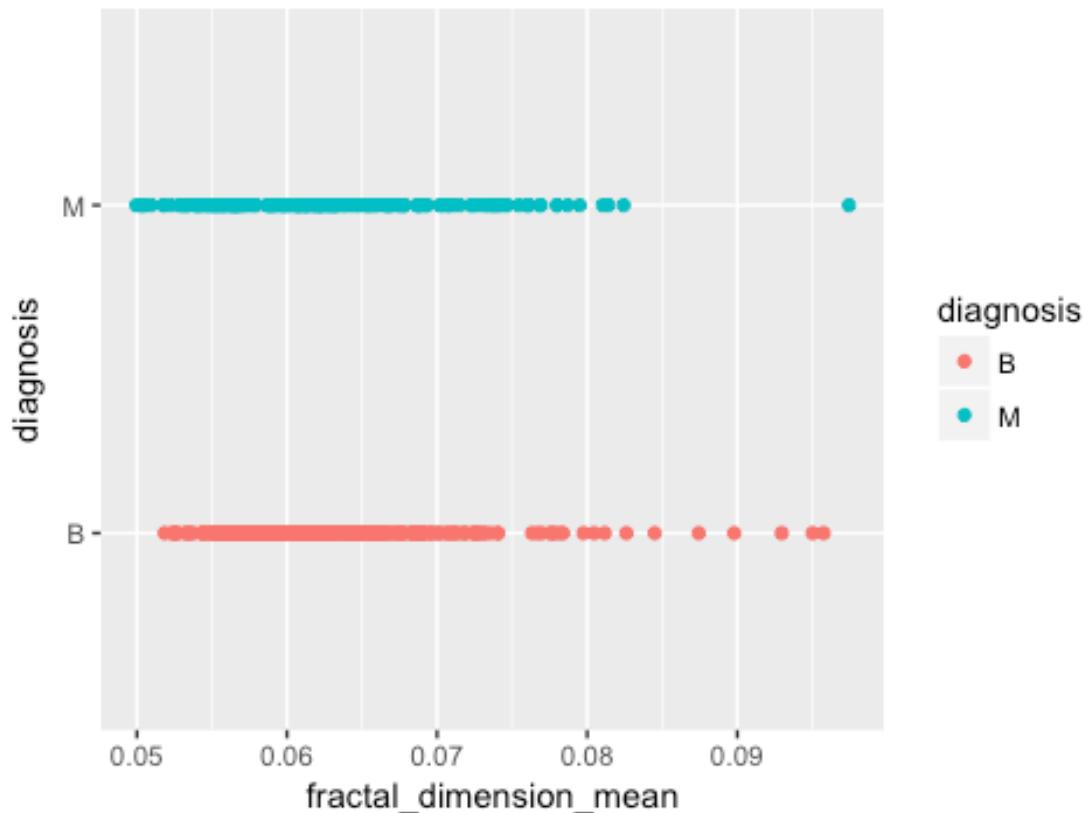
```
#symmetry_mean
sc_symmetry_mean <- ggplot(data = cancer, aes(x = symmetry_mean, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Symmetry Mean")
sc_symmetry_mean
```

Scatterplot Symmetry Mean



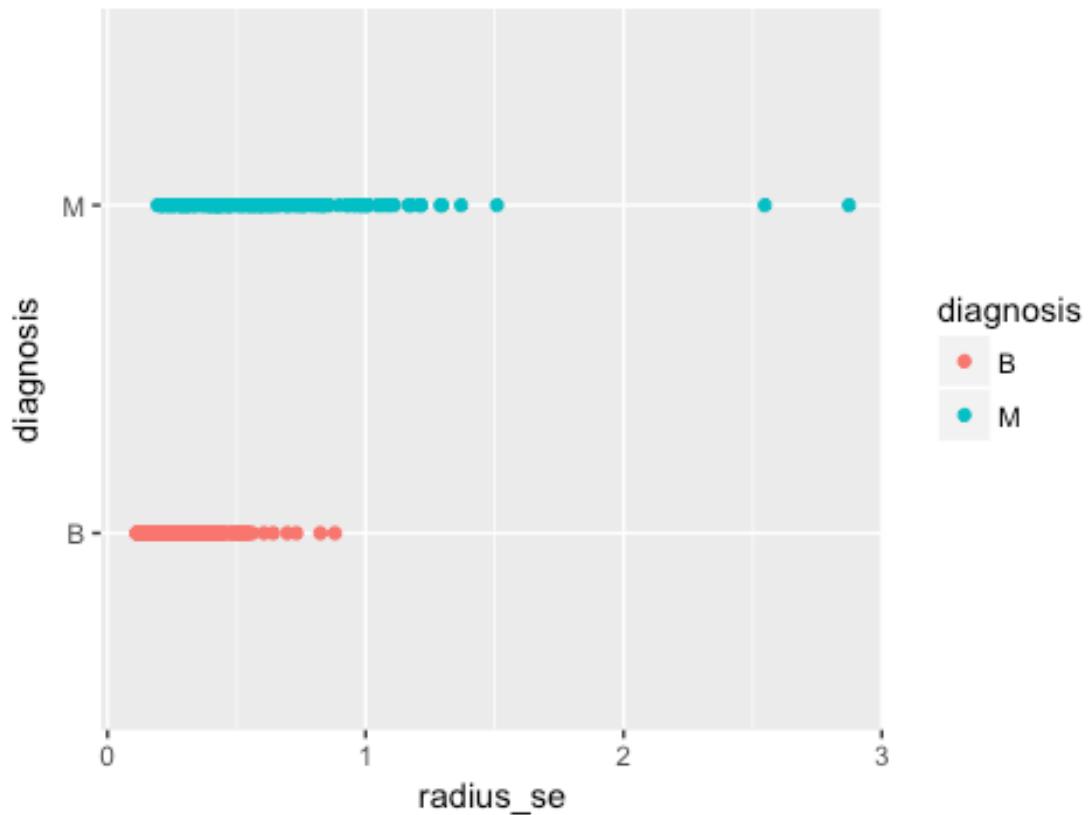
```
#fractaldim_mean
sc_fractaldim_mean <- ggplot(data = cancer, aes(x = fractal_dimension_mean, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Fractal Dimension Mean")
sc_fractaldim_mean
```

Scatterplot Fractal Dimension Mean



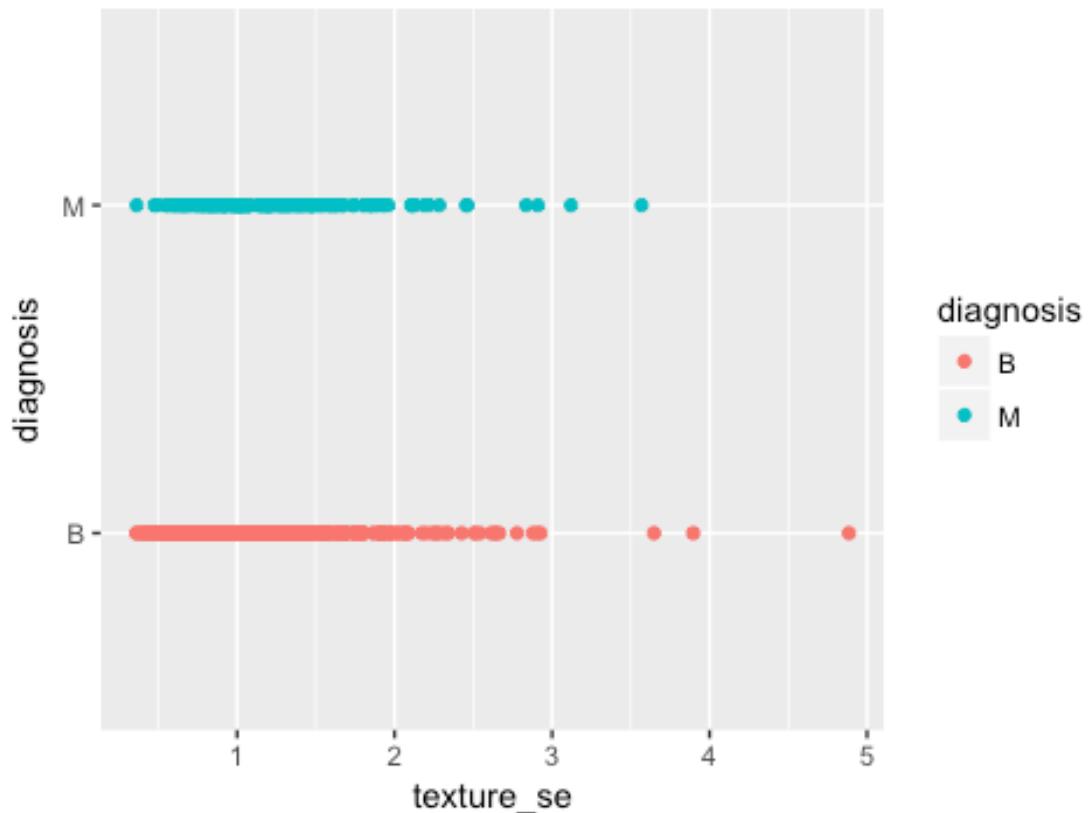
```
##Scatterplots of standard error variables
#radius_se
sc_radius_se <- ggplot(data = cancer, aes(x = radius_se, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Radius Standard Error")
sc_radius_se
```

Scatterplot Radius Standard Error



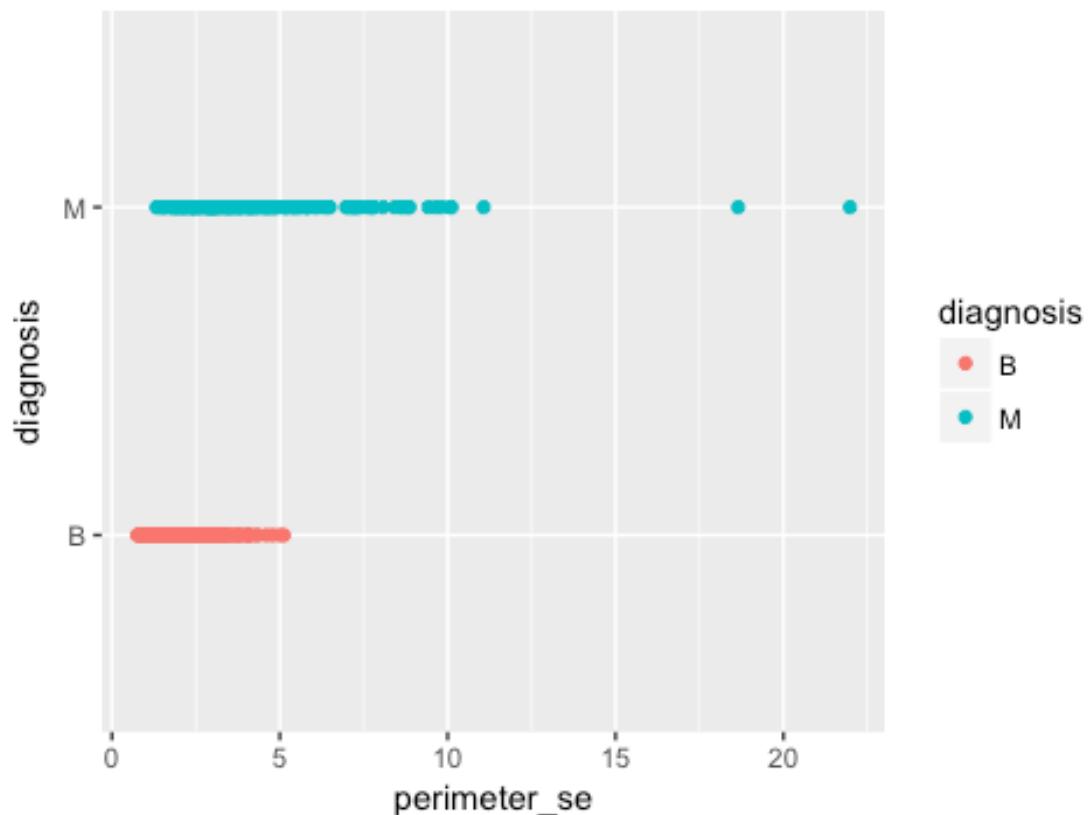
```
#radius_se
sc_texture_se <- ggplot(data = cancer, aes(x = texture_se, y = diagnosis)) +
  aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Texture Standard Error")
sc_texture_se
```

Scatterplot Texture Standard Error



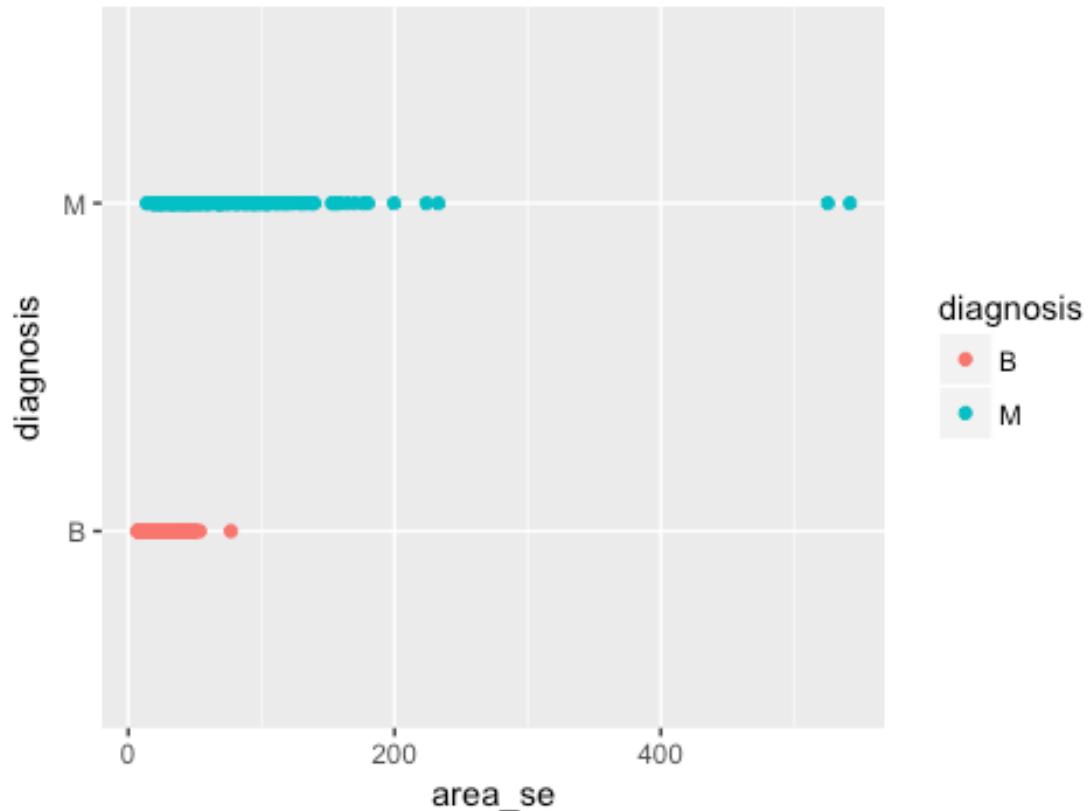
```
#perimeter_se
sc_perimeter_se <- ggplot(data = cancer, aes(x = perimeter_se, y = diagnosis) +
  aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Perimeter Standard Error")
sc_perimeter_se
```

Scatterplot Perimeter Standard Error



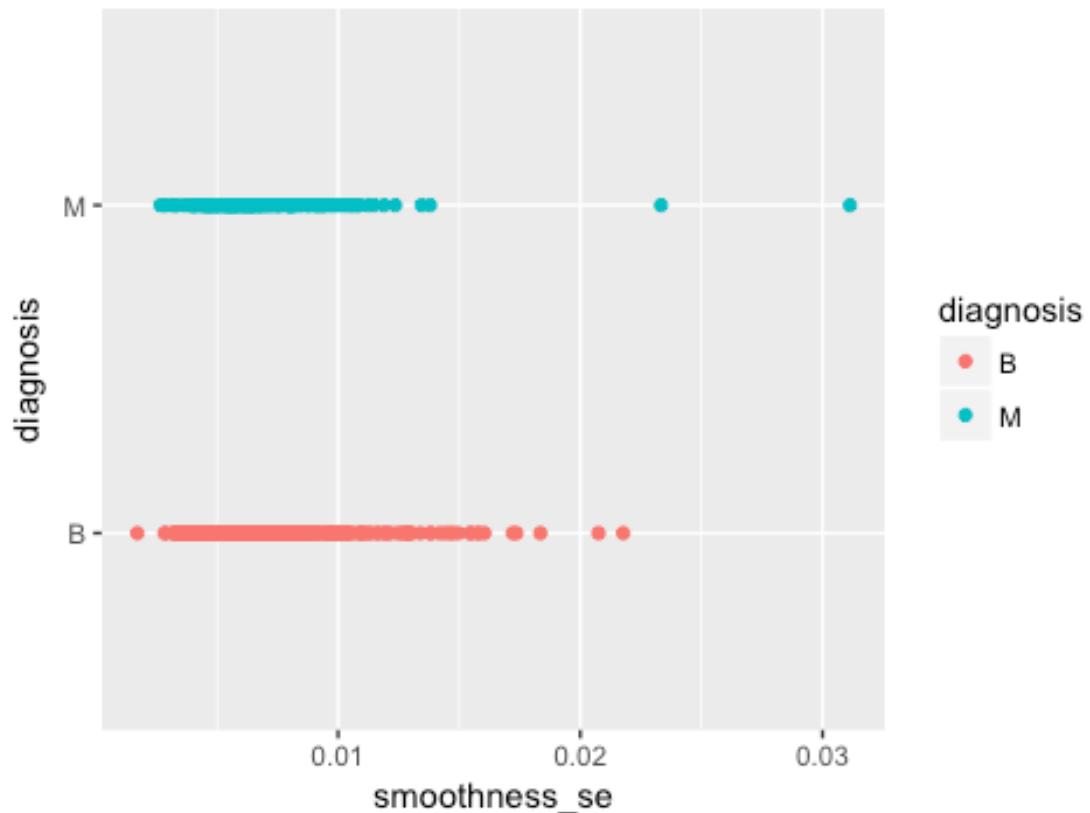
```
#area_se
sc_area_se <- ggplot(data = cancer, aes(x = area_se, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Area Standard Error")
sc_area_se
```

Scatterplot Area Standard Error



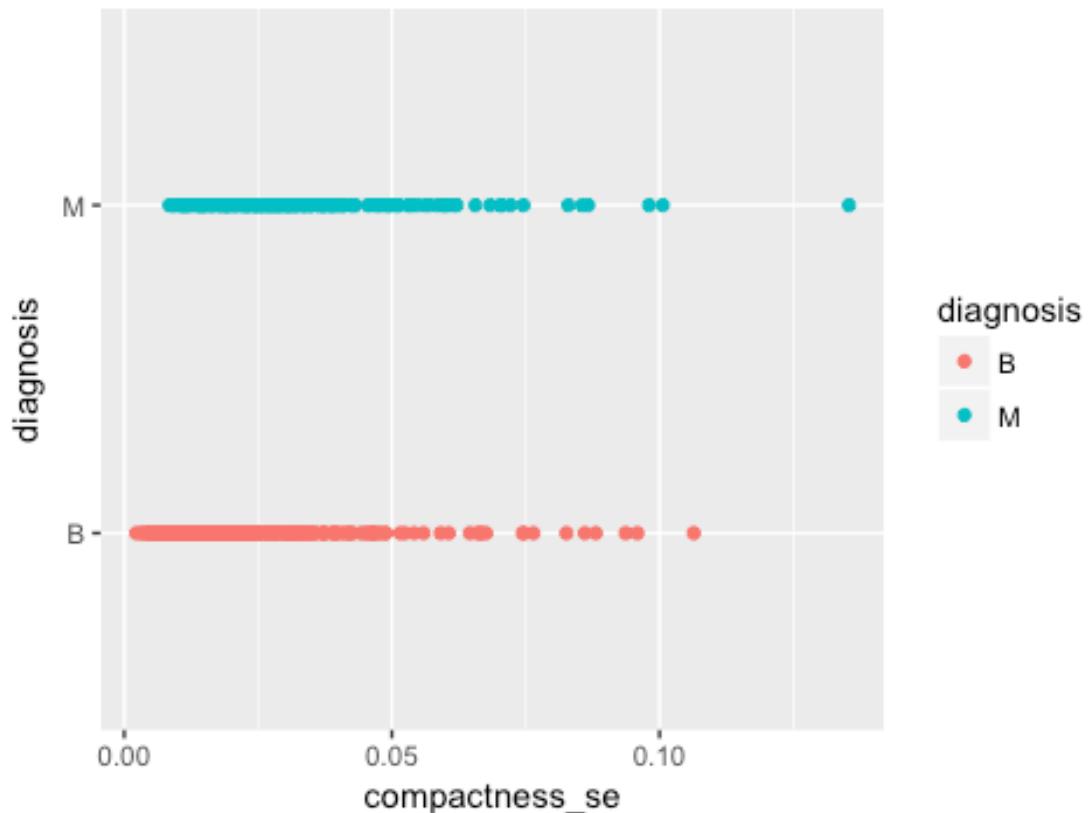
```
#smoothness_se
sc_smoothness_se <- ggplot(data = cancer, aes(x = smoothness_se, y = diagnosis)) +
  aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Smoothness Standard Error")
sc_smoothness_se
```

Scatterplot Smoothness Standard Error



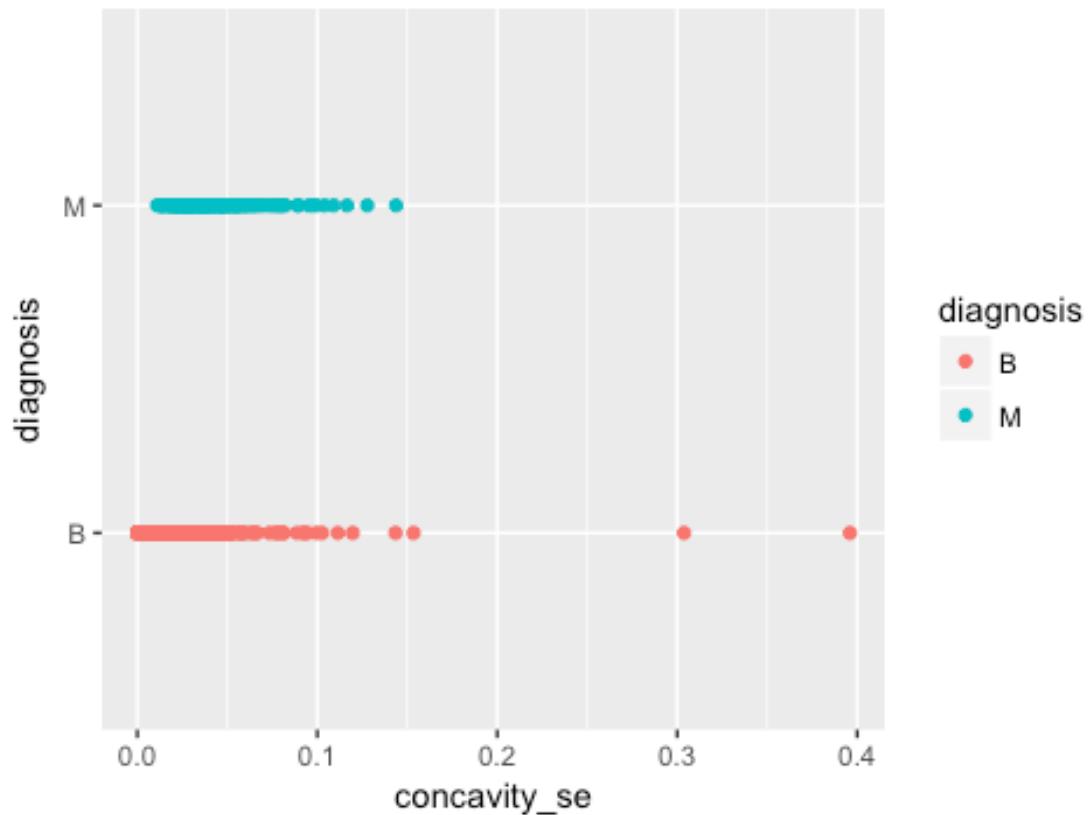
```
#compactness_se
sc_compactness_se <- ggplot(data = cancer, aes(x = compactness_se, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Compactness Standard Error")
sc compactness se
```

Scatterplot Compactness Standard Error



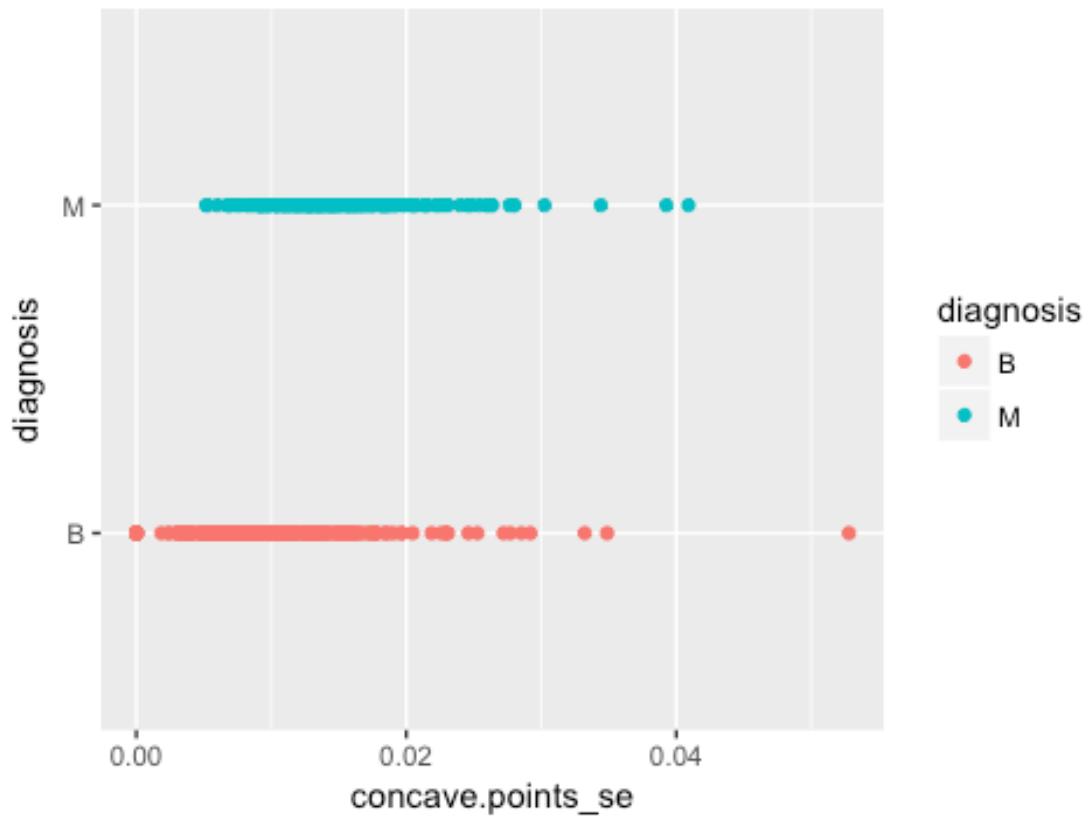
```
#concavity_se
sc_concavity_se <- ggplot(data = cancer, aes(x = concavity_se, y = diagnosis))
+ aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Concavity Standard Error")
sc concavity se
```

Scatterplot Concavity Standard Error



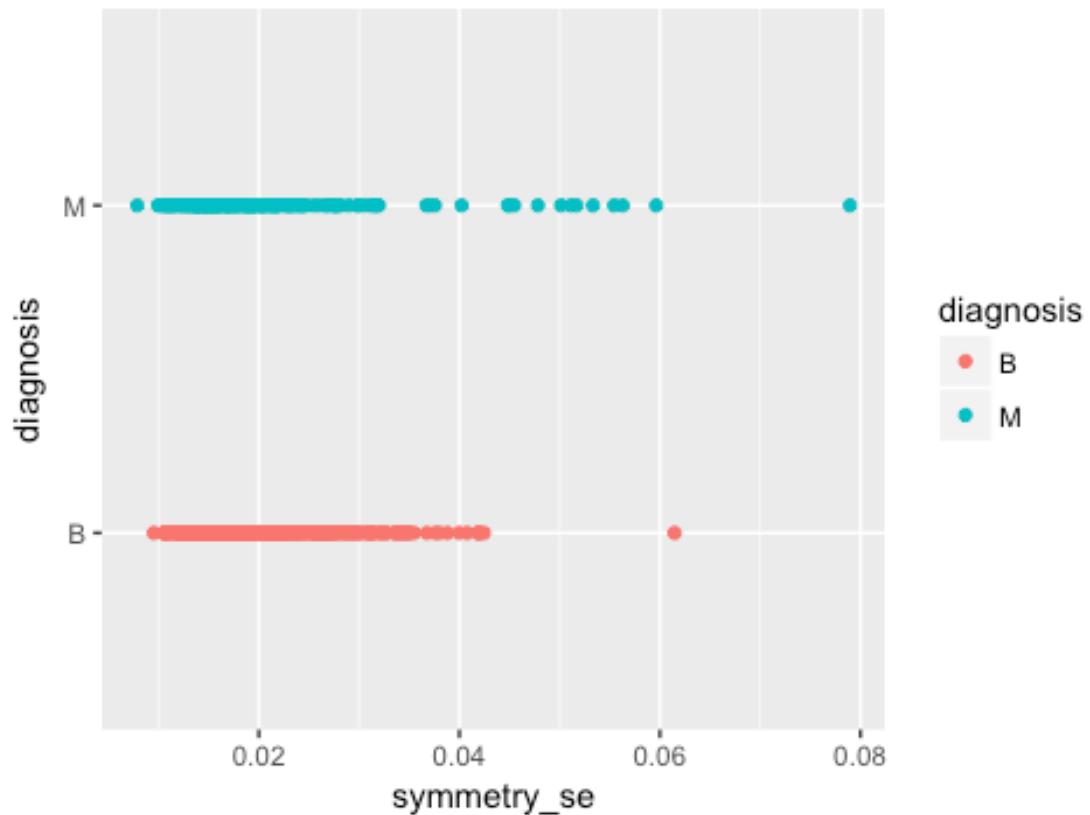
```
#concavepoints_se
sc_concavepoints_se <- ggplot(data = cancer, aes(x = concave.points_se, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Concave Points Standard Error")
sc_concavepoints_se
```

Scatterplot Concave Points Standard Error



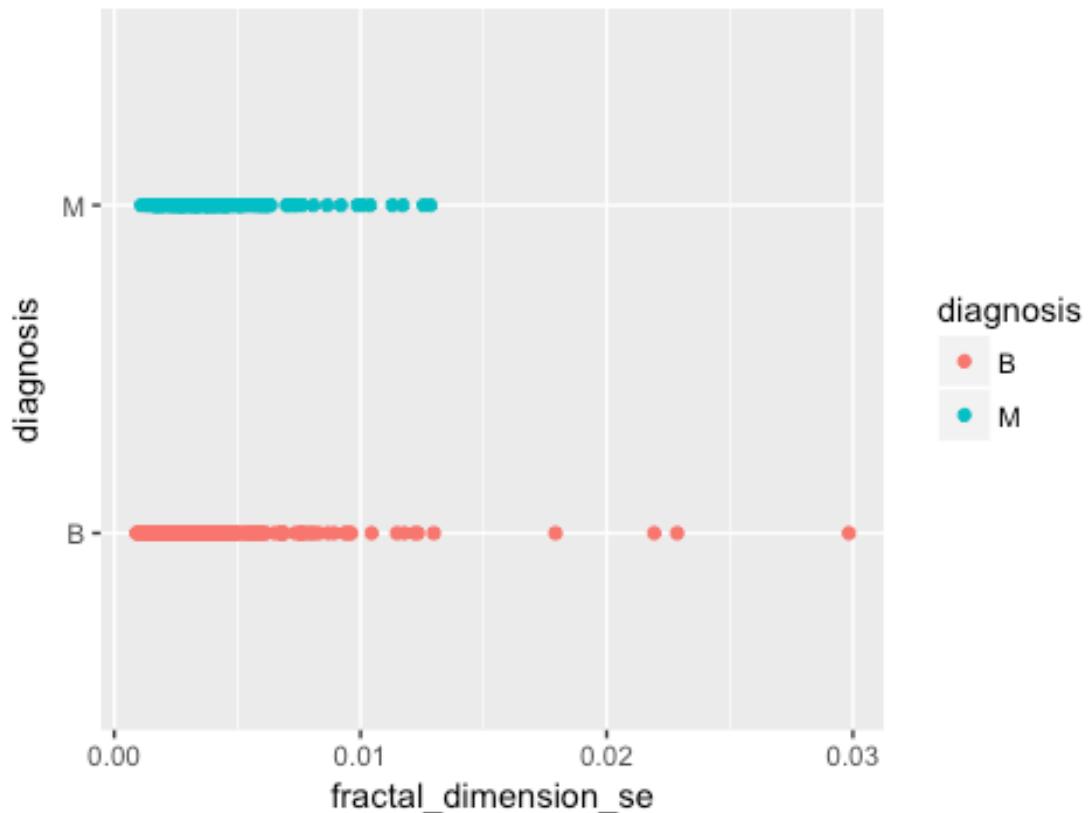
```
#symmetry_se
sc_symmetry_se <- ggplot(data = cancer, aes(x = symmetry_se, y = diagnosis))
+ aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Symmetry Standard Error")
sc_symmetry_se
```

Scatterplot Symmetry Standard Error



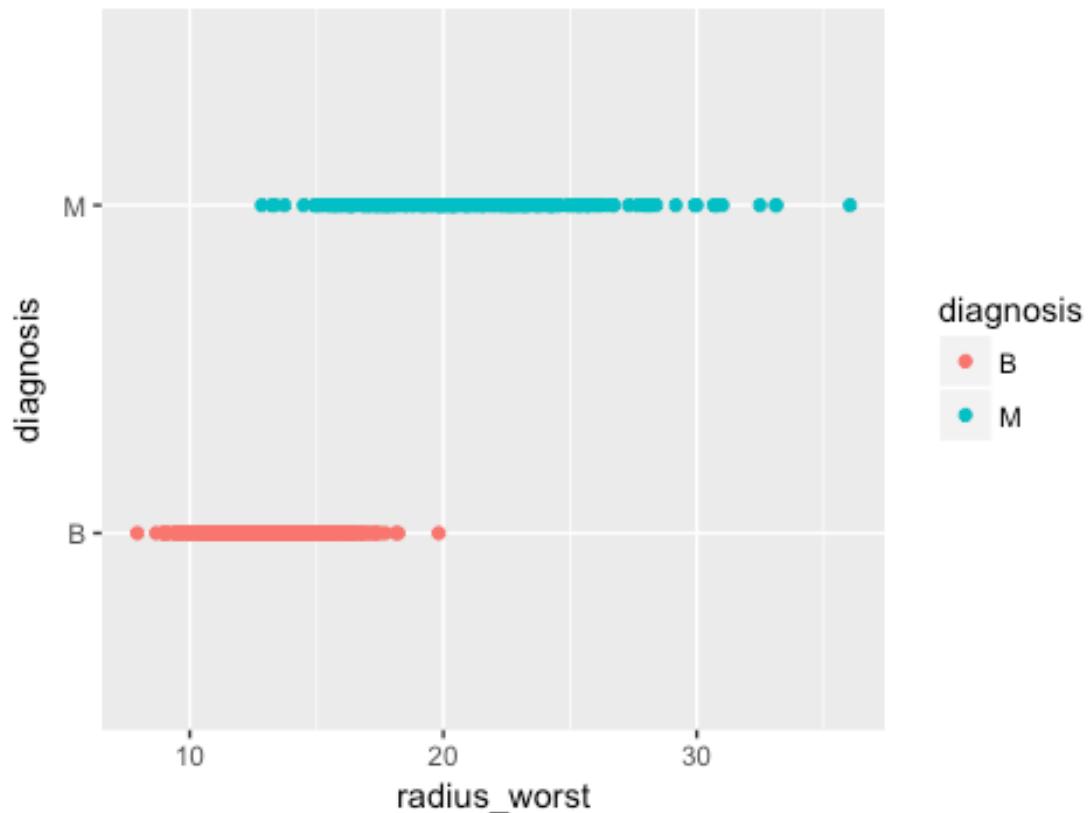
```
#fractaldim_se
sc_fractaldim_se <- ggplot(data = cancer, aes(x = fractal_dimension_se, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Fractal Dimension Standard Error")
sc_fractaldim_se
```

Scatterplot Fractal Dimension Standard Error



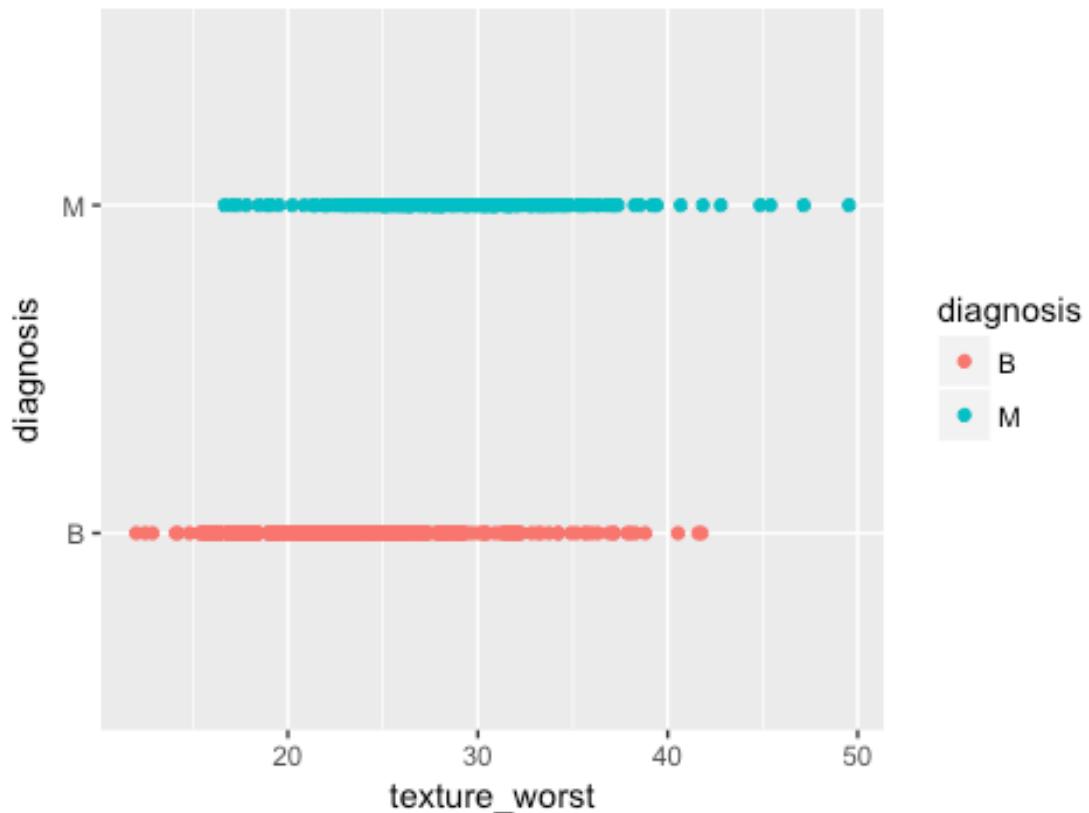
```
##Scatterplots of worst variables
#radius_worst
sc_radius_worst <- ggplot(data = cancer, aes(x = radius_worst, y = diagnosis))
+ aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Radius Maximum")
sc_radius_worst
```

Scatterplot Radius Maximum



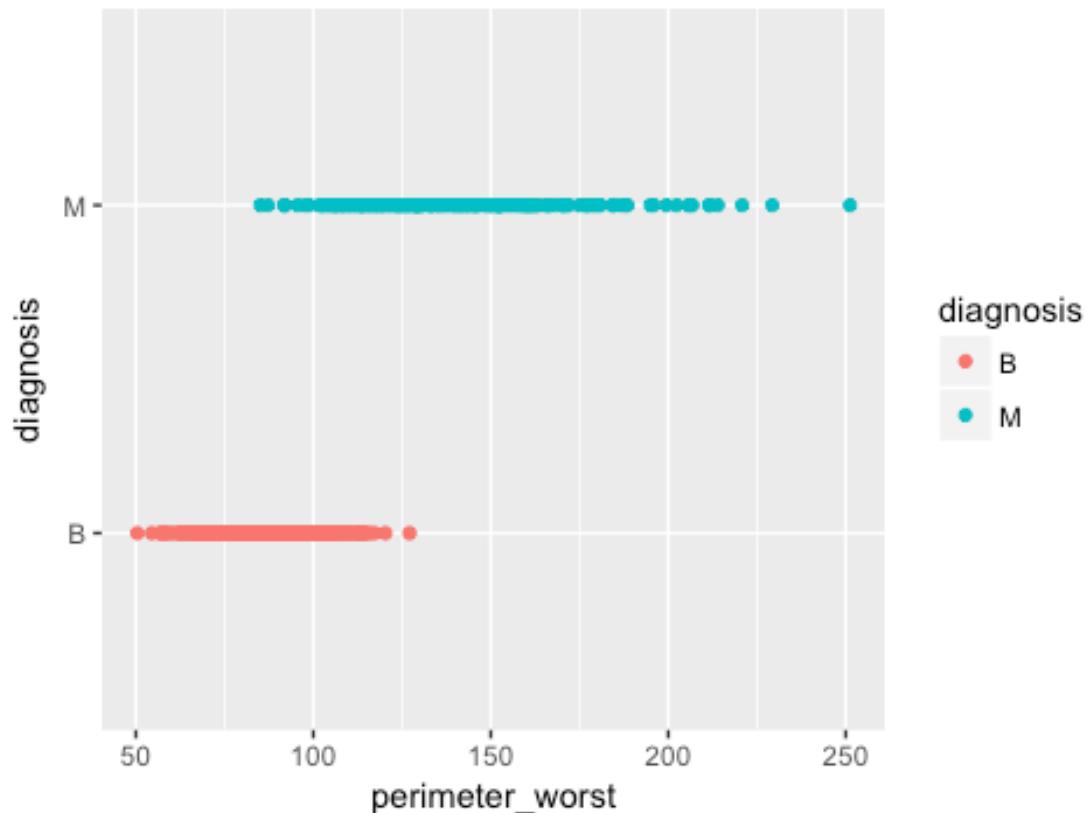
```
#texture_worst
sc_texture_worst <- ggplot(data = cancer, aes(x = texture_worst, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Texture Maximum")
sc_texture_worst
```

Scatterplot Texture Maximum



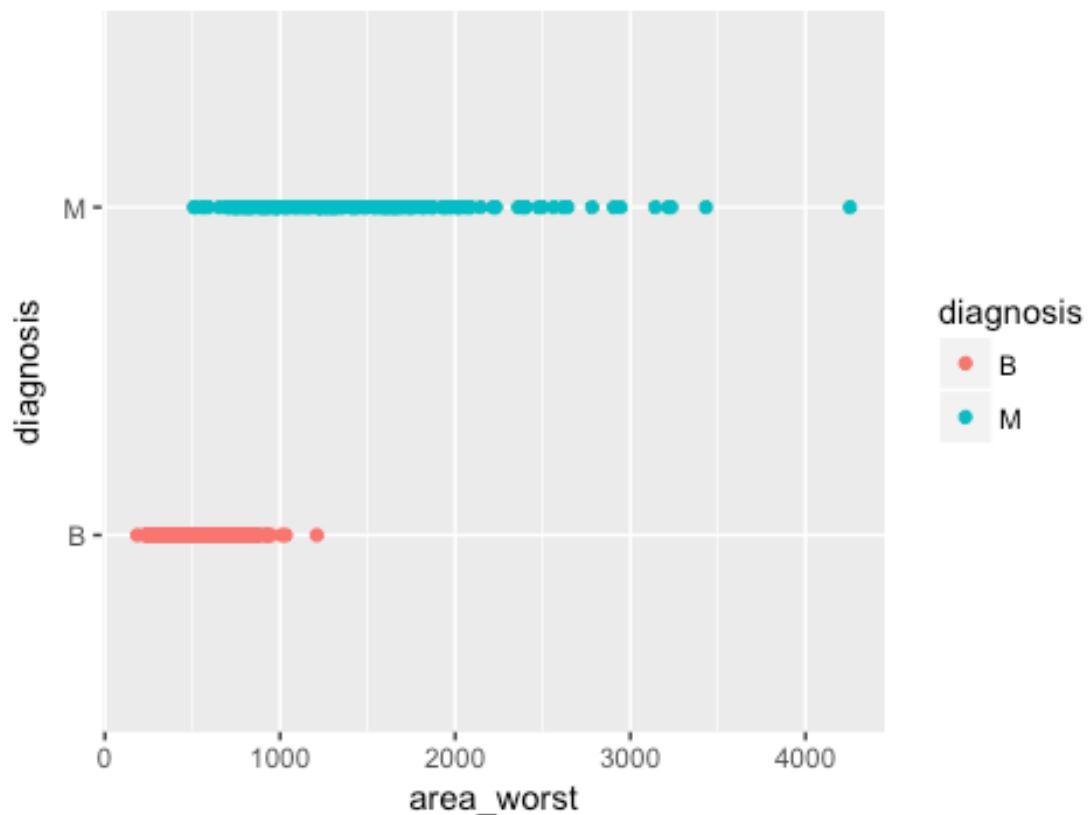
```
#perimeter_worst
sc_perimeter_worst <- ggplot(data = cancer, aes(x = perimeter_worst, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Perimeter Maximum")
sc_perimeter_worst
```

Scatterplot Perimeter Maximum



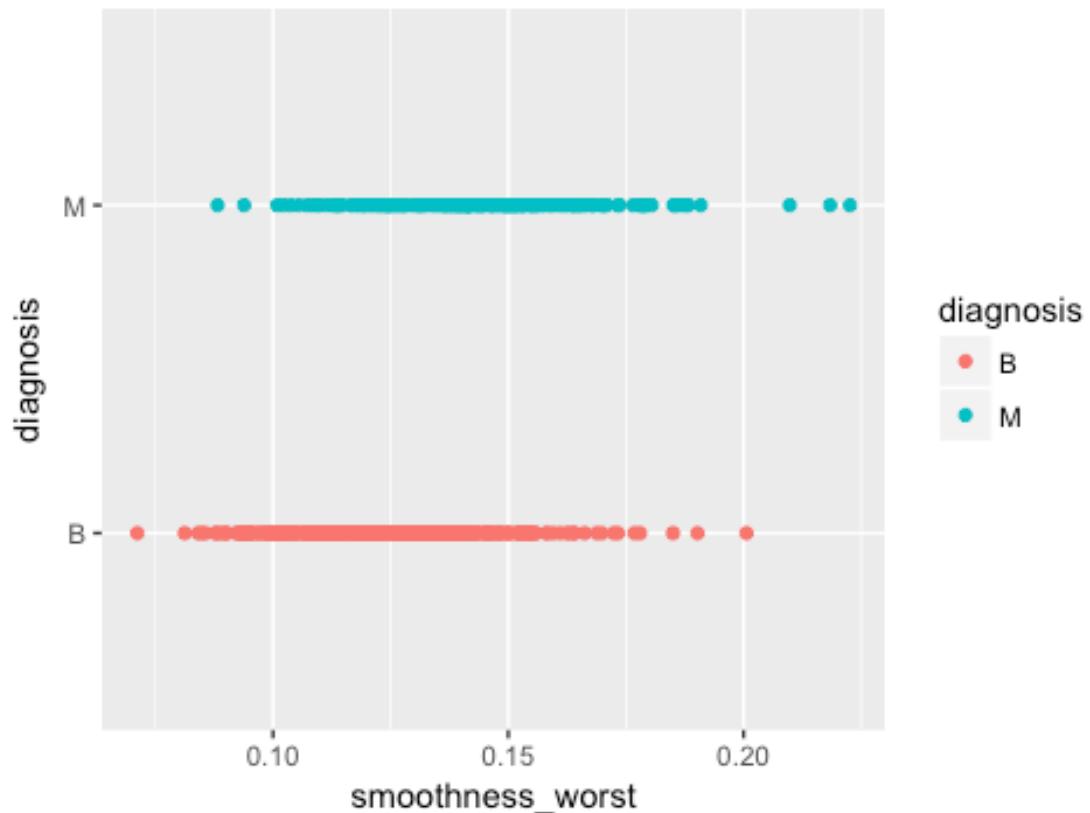
```
#area_worst
sc_area_worst <- ggplot(data = cancer, aes(x = area_worst, y = diagnosis)) +
  aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Area Maximum")
sc_area_worst
```

Scatterplot Area Maximum



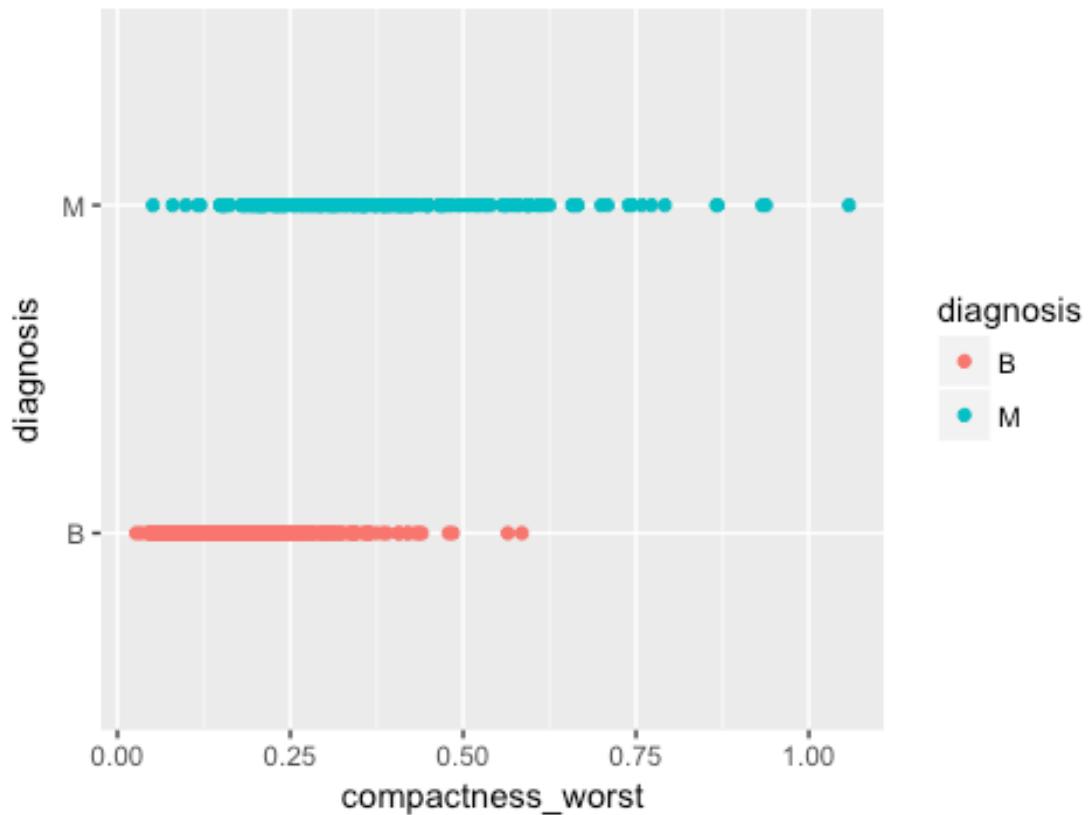
```
#smoothness_worst
sc_smoothness_worst <- ggplot(data = cancer, aes(x = smoothness_worst, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Smoothness Maximum")
sc_smoothness_worst
```

Scatterplot Smoothness Maximum



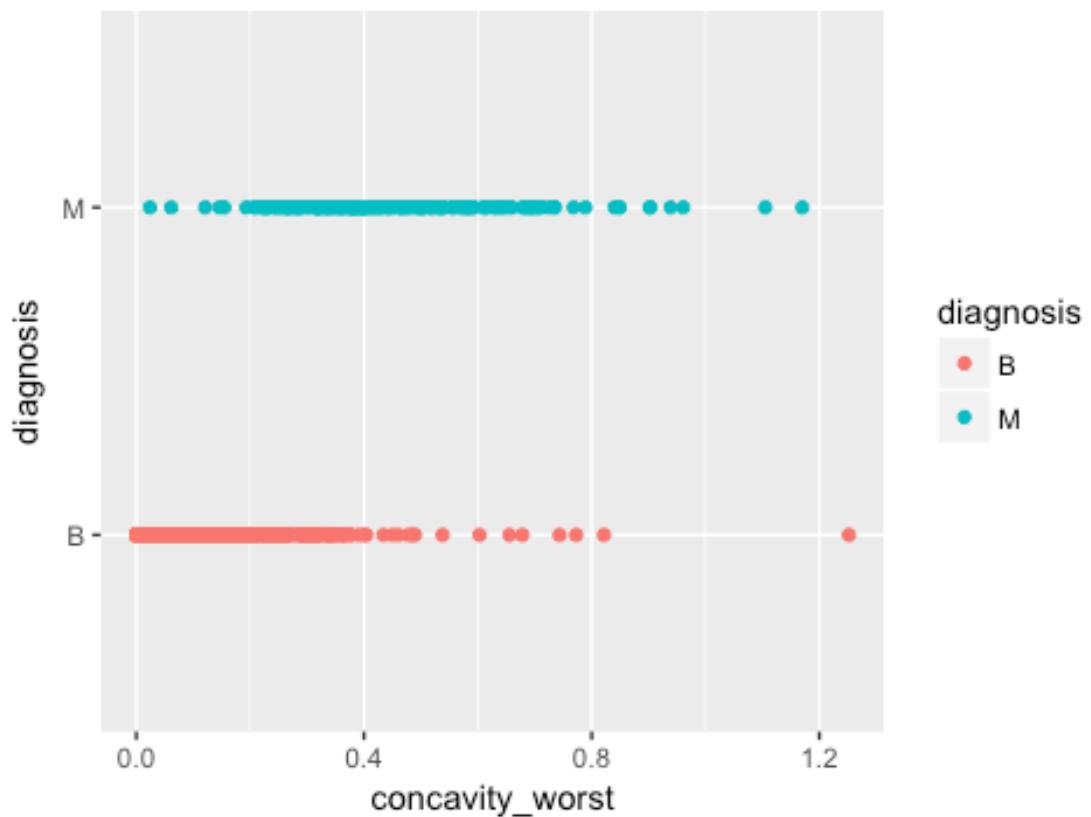
```
#compactness_worst
sc_compactness_worst <- ggplot(data = cancer, aes(x = compactness_worst, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Compactness Maximum")
sc_compactness_worst
```

Scatterplot Compactness Maximum



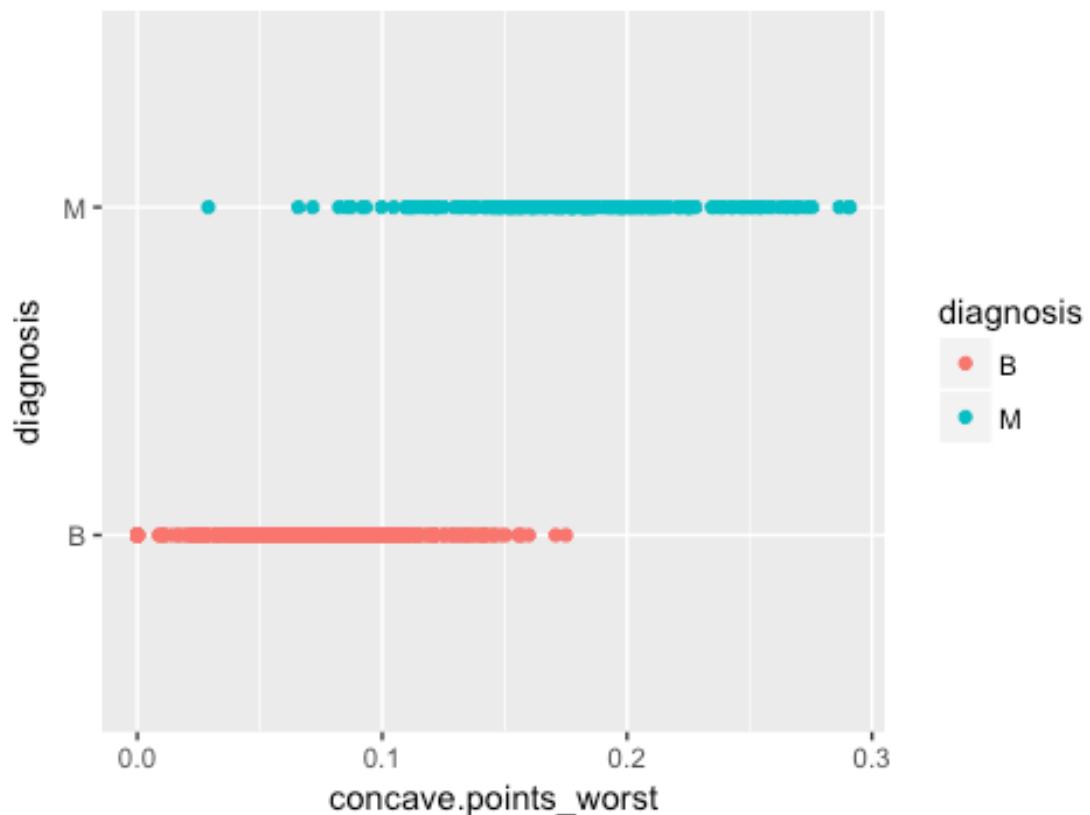
```
#concavity_worst
sc_concavity_worst <- ggplot(data = cancer, aes(x = concavity_worst, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Concavity Maximum")
sc_concavity_worst
```

Scatterplot Concavity Maximum



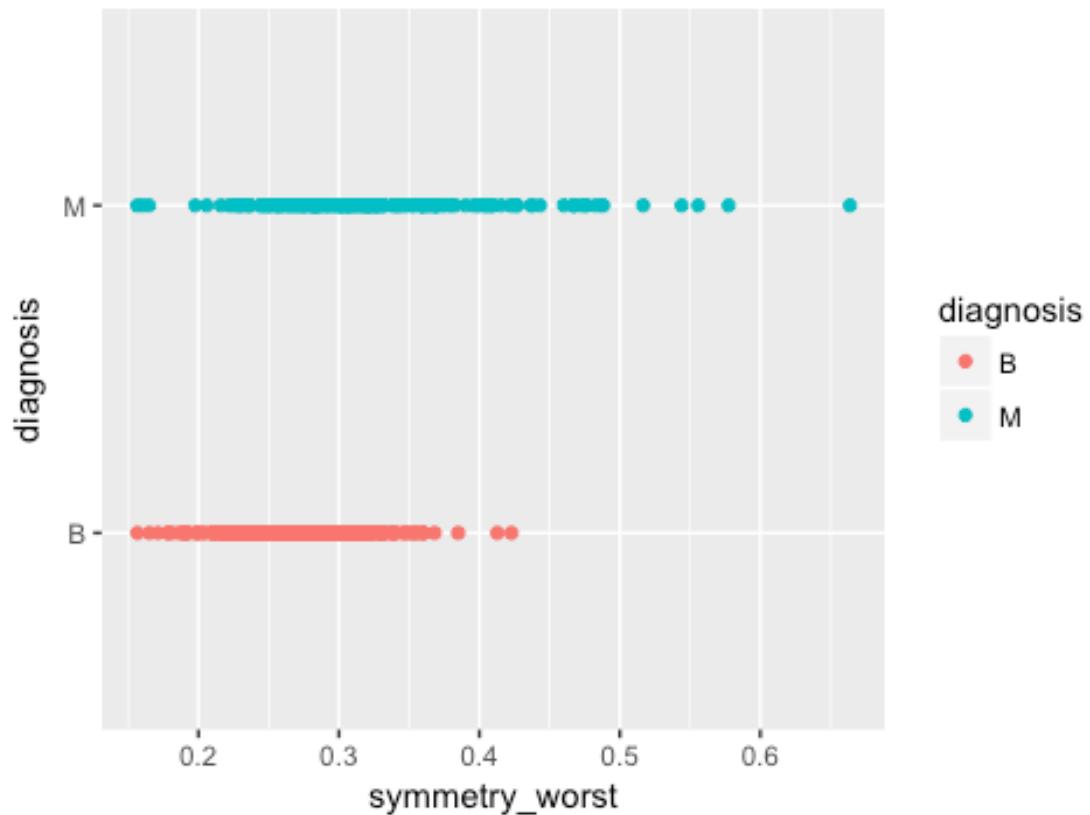
```
#concavepoints_worst
sc_concavepoints_worst <- ggplot(data = cancer, aes(x = concave.points_worst,
y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Concave Points Maximum")
sc_concavepoints_worst
```

Scatterplot Concave Points Maximum



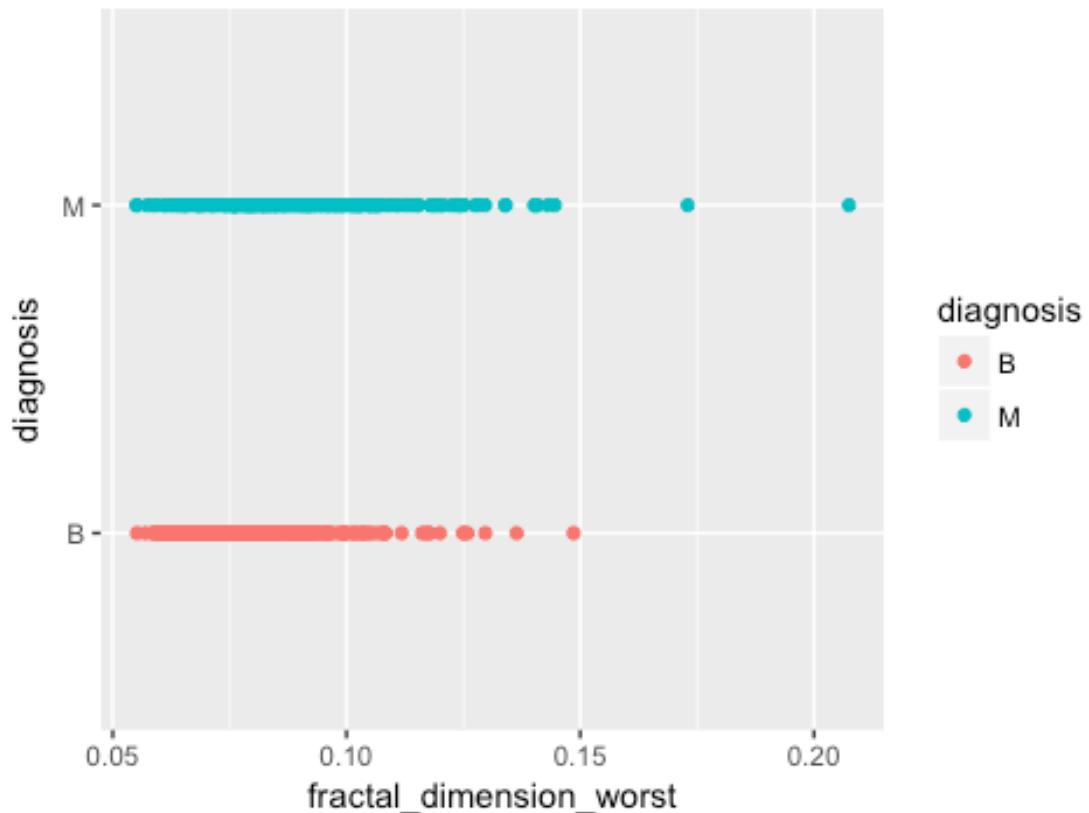
```
#symmetry_worst
sc_symmetry_worst <- ggplot(data = cancer, aes(x = symmetry_worst, y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Symmetry Maximum")
sc_symmetry_worst
```

Scatterplot Symmetry Maximum



```
#fractaldim_worst
sc_fractaldim_worst <- ggplot(data = cancer, aes(x = fractal_dimension_worst,
y = diagnosis)) + aes(color=diagnosis) +
  geom_point() + labs(title = "Scatterplot Fractal Dimension Maximum")
sc_fractaldim_worst
```

Scatterplot Fractal Dimension Maximum

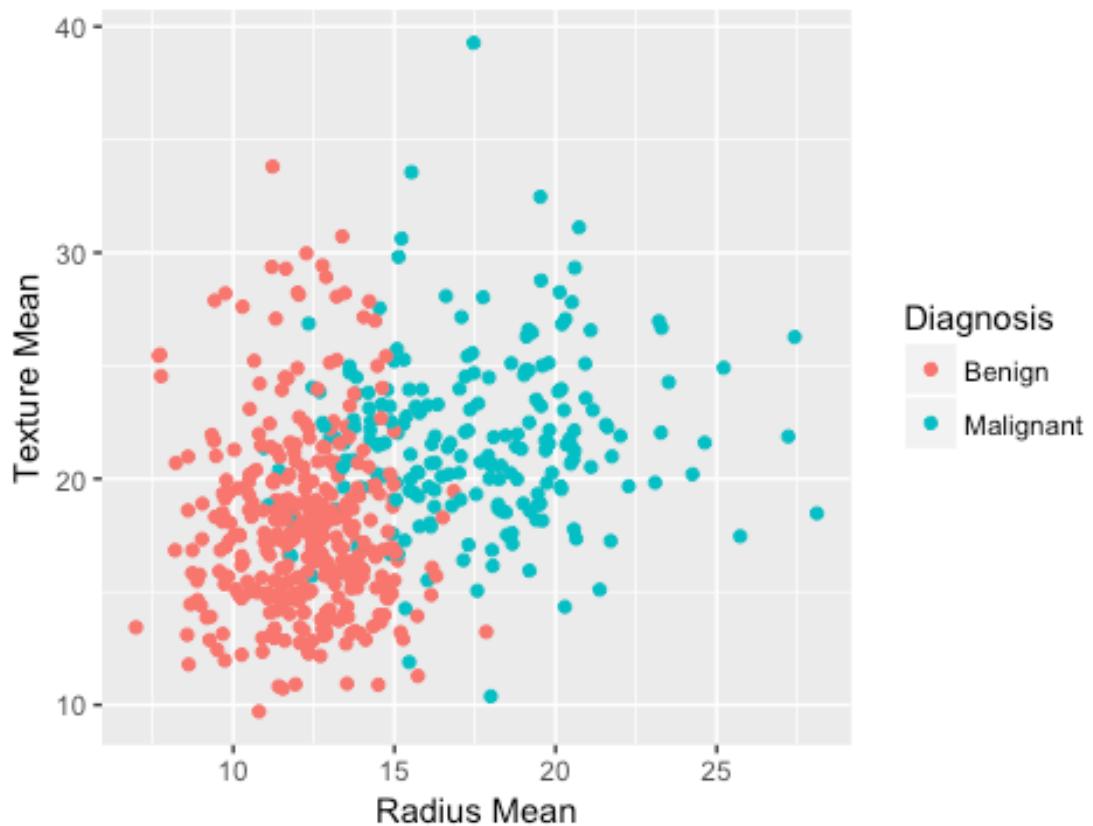


```
# Multivariate EDA

#Mean variable relationships

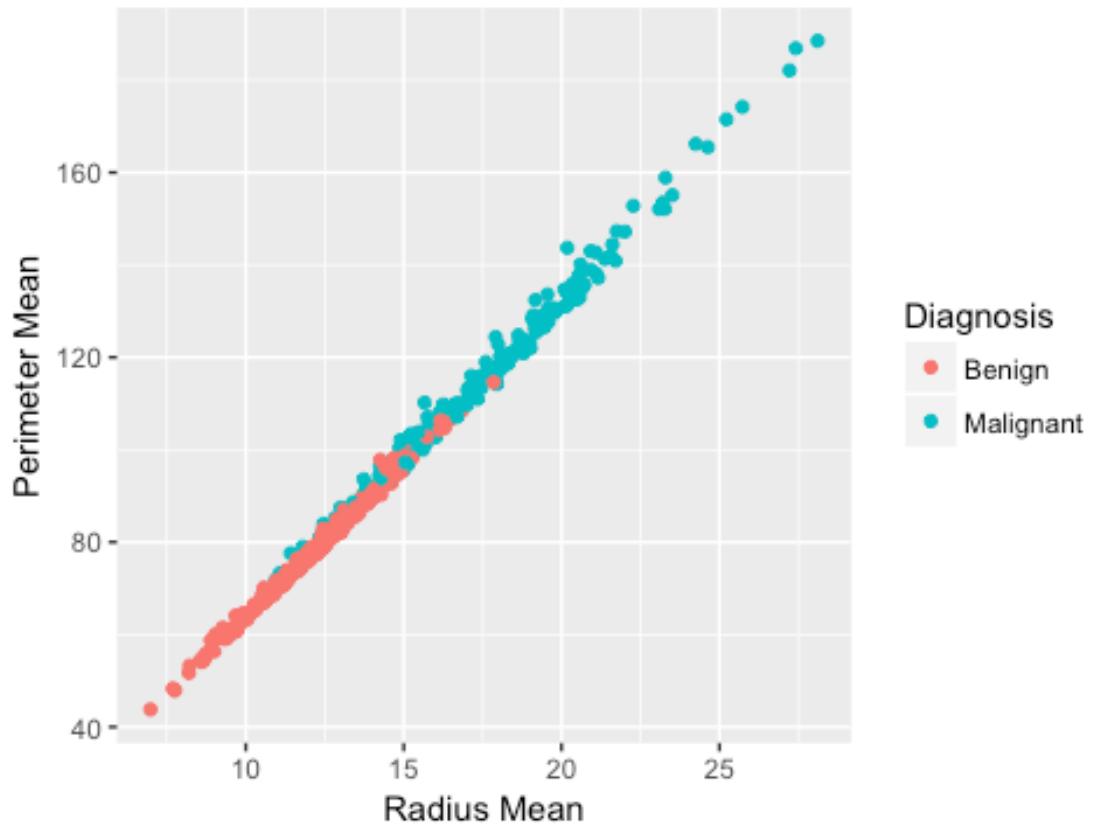
#texture_mean vs. radius_mean
radius_texture_mean_scatter <- ggplot(data = cancer, aes(x = radius_mean, y = texture_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Mean", y = "Texture Mean", title = "Texture Mean vs. Radius Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius texture mean scatter
```

Texture Mean vs. Radius Mean

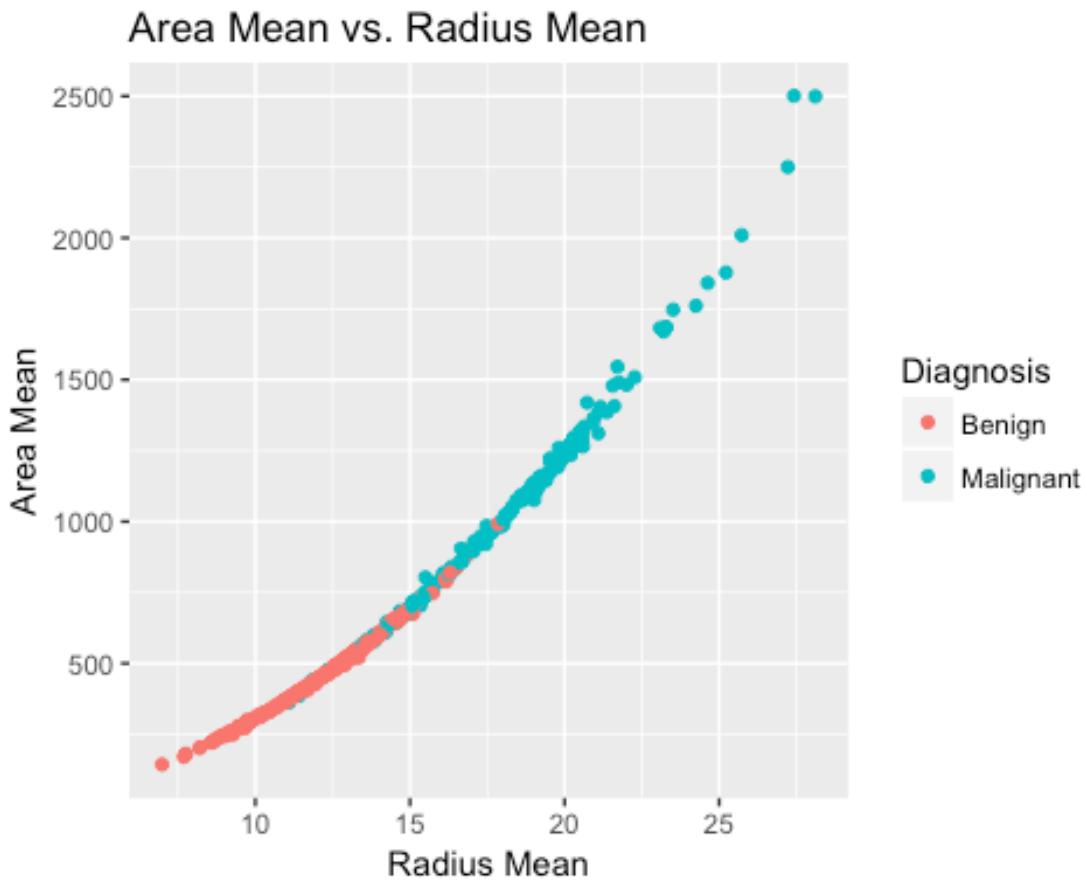


```
#perimeter_mean vs. radius_mean
radius_perimeter_mean_scatter <- ggplot(data = cancer, aes(x = radius_mean, y = perimeter_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Mean", y = "Perimeter Mean", title = "Perimeter Mean vs. Radius Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_perimeter_mean_scatter
```

Perimeter Mean vs. Radius Mean

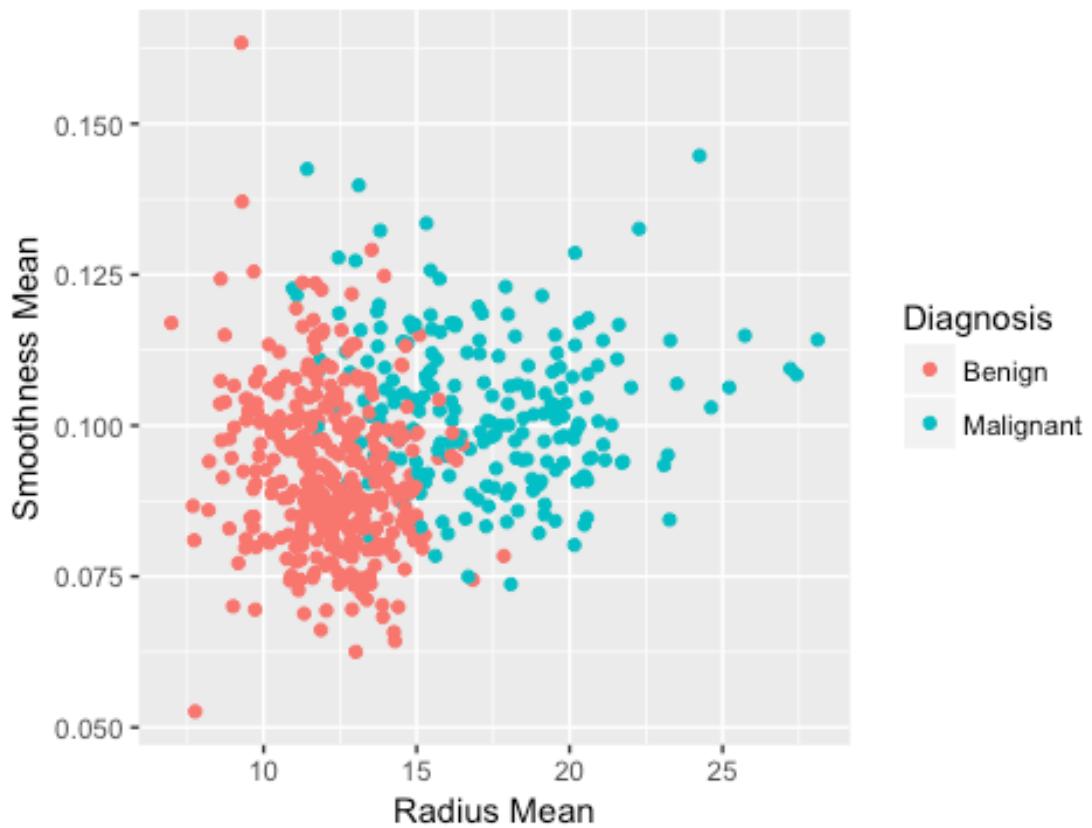


```
#area_mean vs. radius_mean
radius_area_mean_scatter <- ggplot(data = cancer, aes(x = radius_mean, y = area_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Mean", y = "Area Mean", title = "Area Mean vs. Radius Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_area_mean_scatter
```



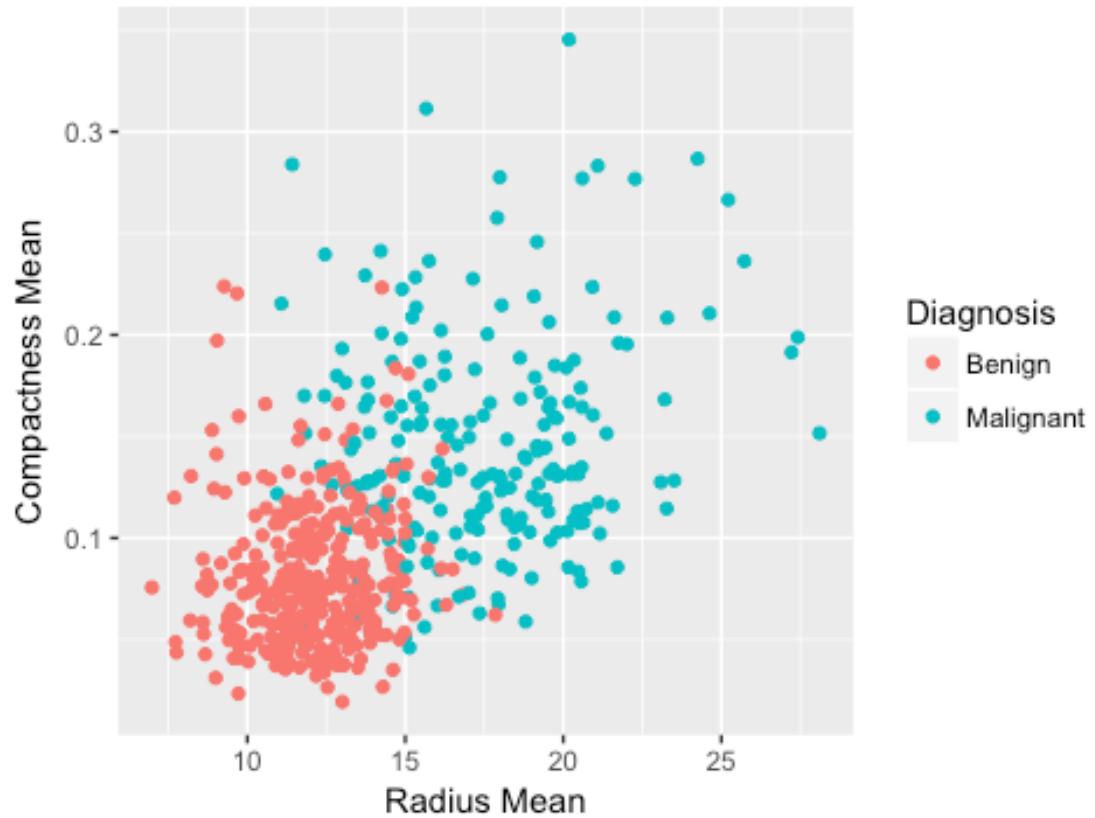
```
#smoothness_mean vs. radius_mean
radius_smoothness_mean_scatter <- ggplot(data = cancer, aes(x = radius_mean,
y = smoothness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Mean", y = "Smoothness Mean", title = "Smoothness Mean vs. Radius Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_smoothness_mean_scatter
```

Smoothness Mean vs. Radius Mean



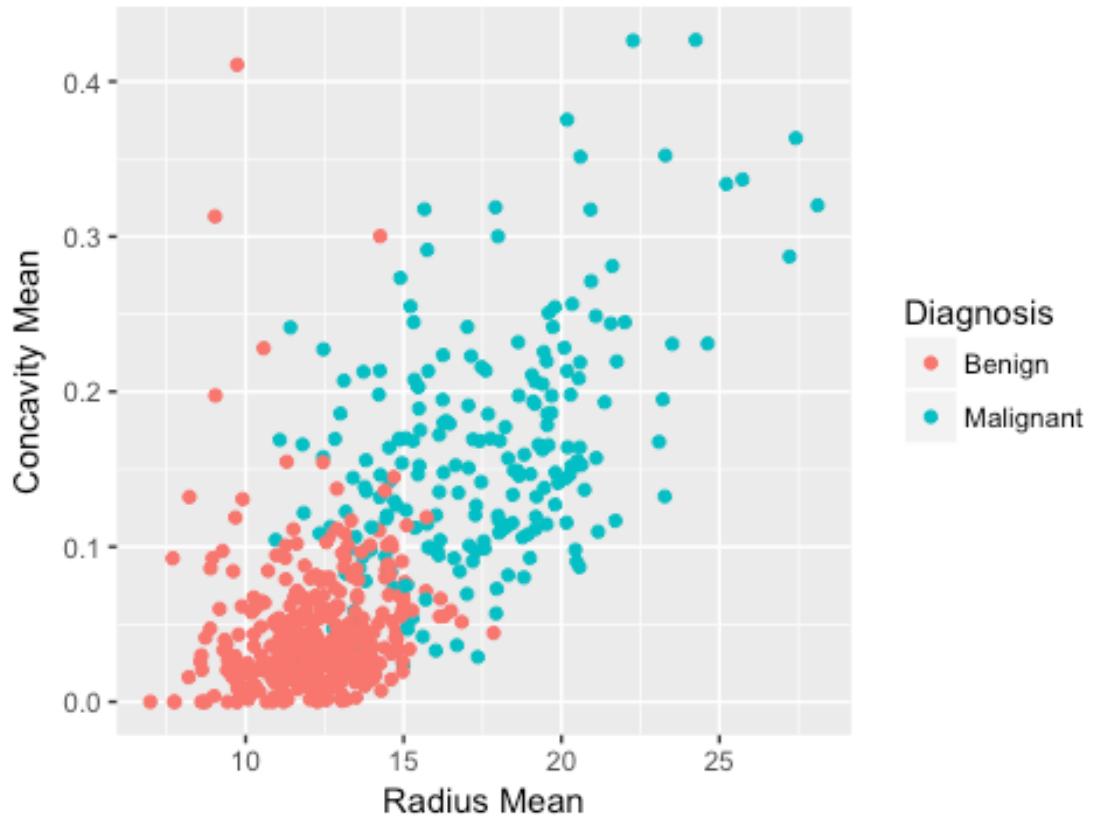
```
#compactness_mean vs. radius_mean
radius_compactness_mean_scatter <- ggplot(data = cancer, aes(x = radius_mean,
y = compactness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Mean", y = "Compactness Mean", title = "Compactness Mean vs. Radius Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_compactness_mean_scatter
```

Compactness Mean vs. Radius Mean

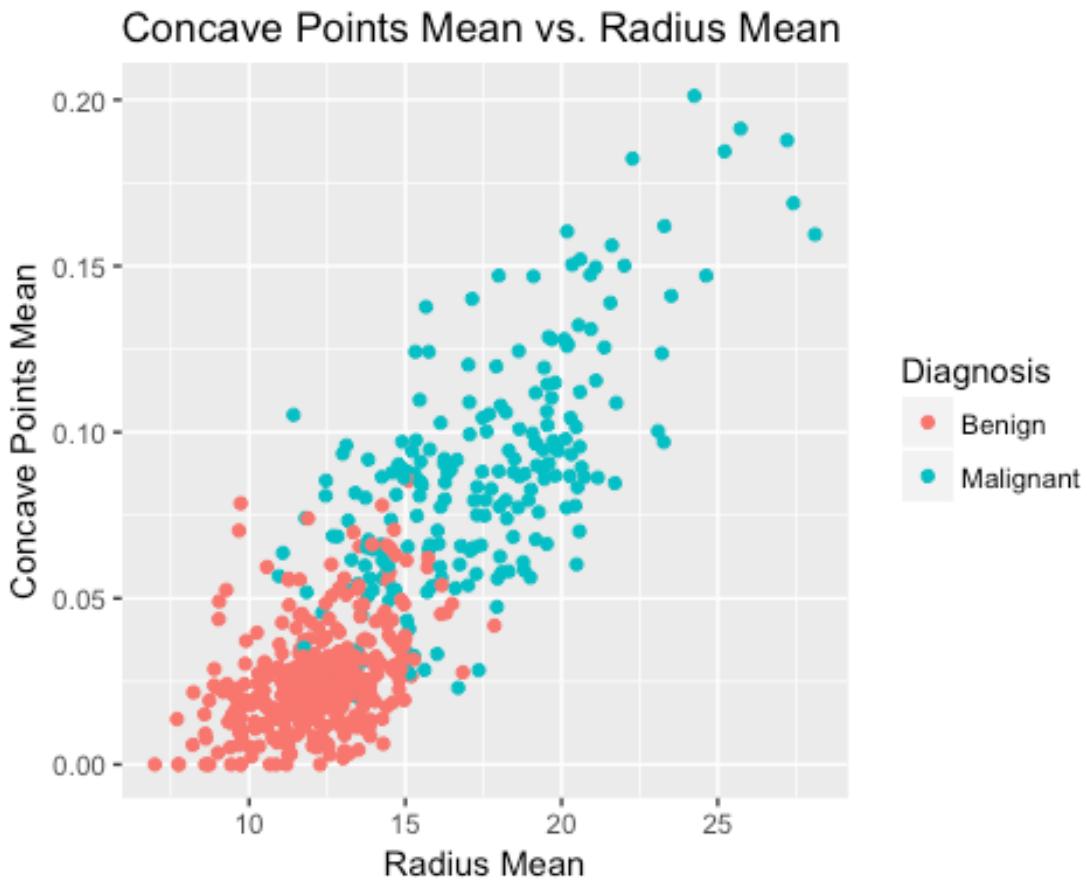


```
#concavity_mean vs. radius_mean
radius_concavity_mean_scatter <- ggplot(data = cancer, aes(x = radius_mean, y = concavity_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Mean", y = "Concavity Mean", title = "Concavity Mean vs. Radius Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_concavity_mean_scatter
```

Concavity Mean vs. Radius Mean

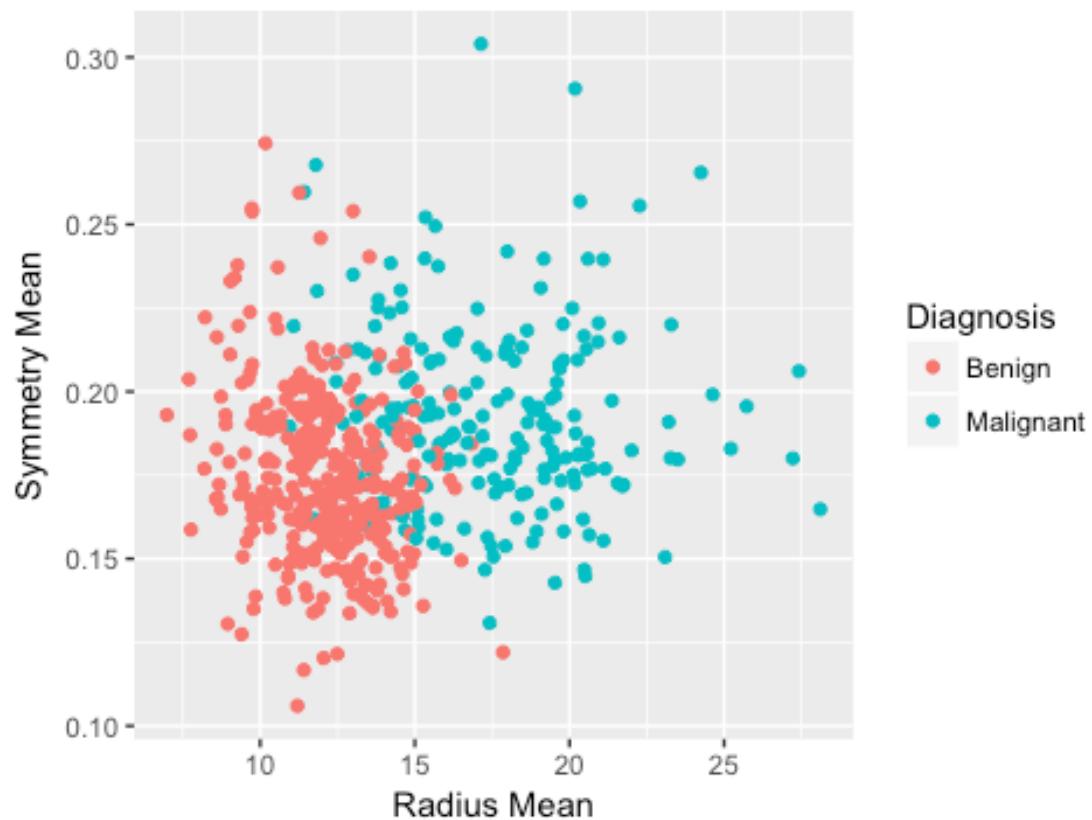


```
#concave.points_mean vs. radius_mean
radius_concave.points_mean_scatter <- ggplot(data = cancer, aes(x = radius_mean, y = concave.points_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Mean", y = "Concave Points Mean", title = "Concave Points Mean vs. Radius Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_concave.points_mean_scatter
```



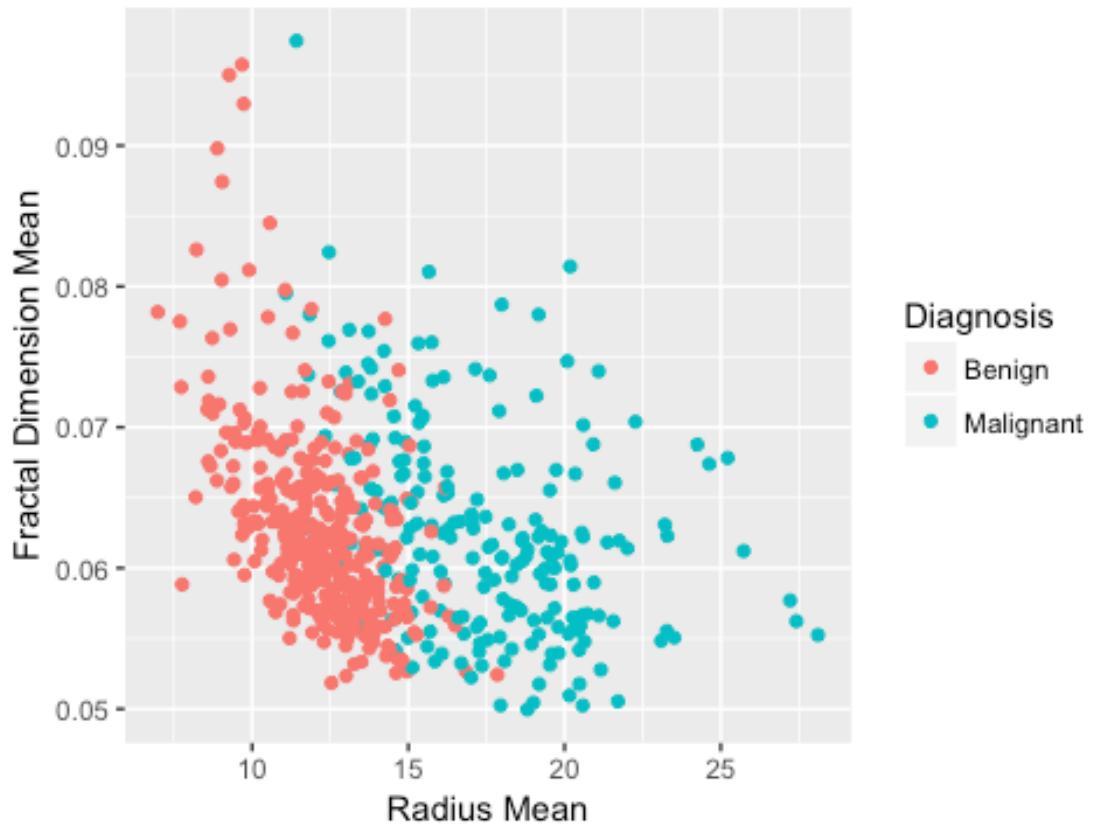
```
#symmetry_mean vs. radius_mean
radius_symmetry_mean_scatter <- ggplot(data = cancer, aes(x = radius_mean, y = symmetry_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Mean", y = "Symmetry Mean", title = "Symmetry Mean vs. Radius Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_symmetry_mean_scatter
```

Symmetry Mean vs. Radius Mean



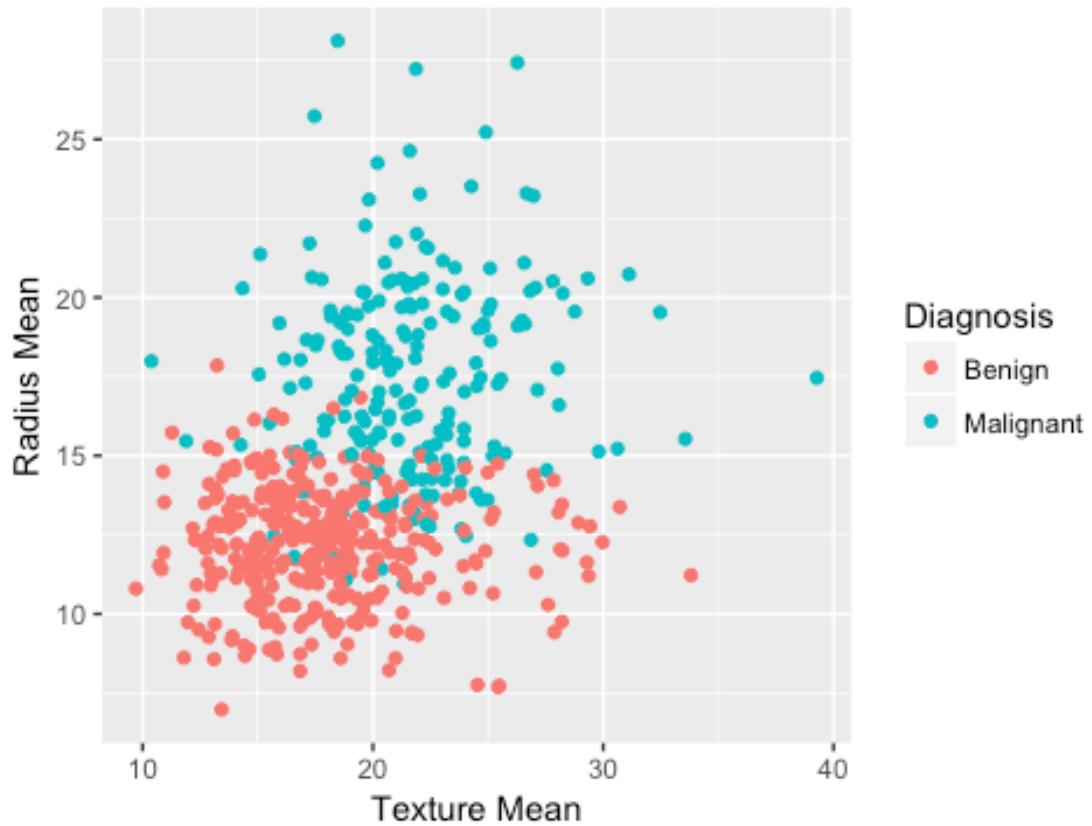
```
#fractal_dimension_mean vs. radius_mean
radius_fractal_dimension_mean_scatter <- ggplot(data = cancer, aes(x = radius_mean, y = fractal_dimension_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Mean", y = "Fractal Dimension Mean", title =
  "Fractal Dimension Mean vs. Radius Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_fractal_dimension_mean_scatter
```

Fractal Dimension Mean vs. Radius Mean



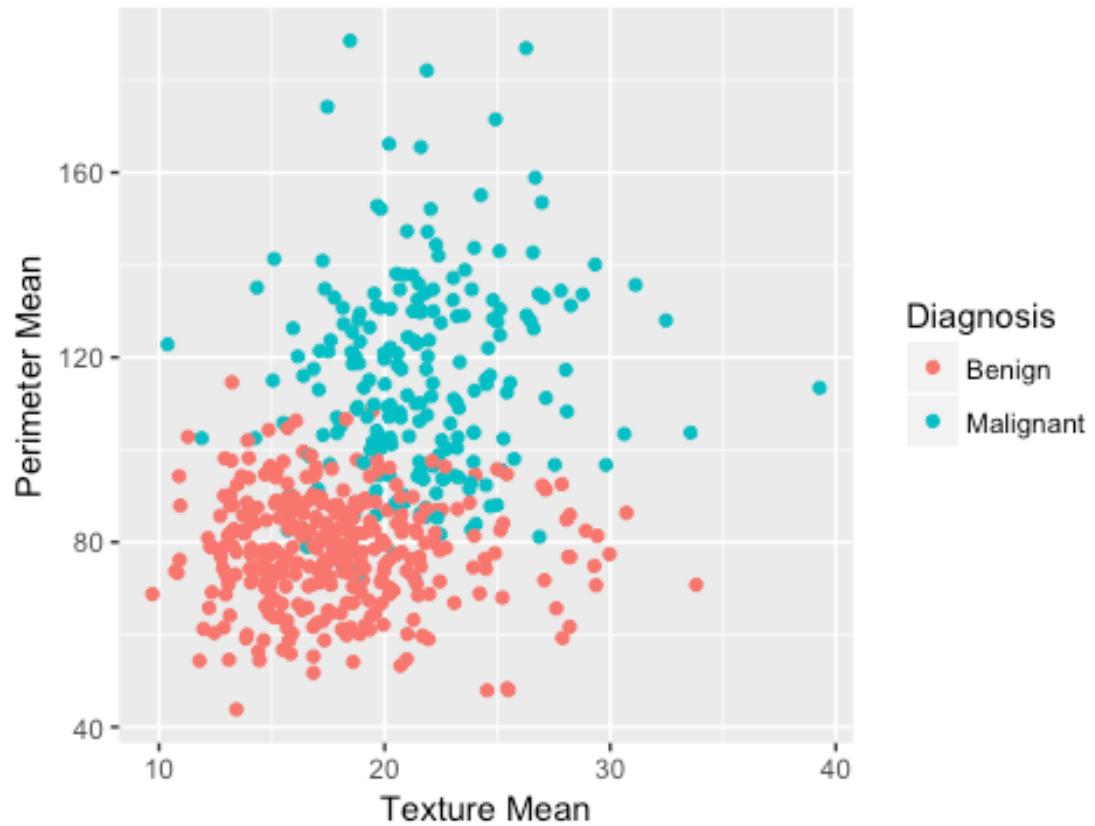
```
#radius_mean vs. texture_mean
texture_radius_mean_scatter <- ggplot(data = cancer, aes(x = texture_mean, y = radius_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Mean", y = "Radius Mean", title = "Radius Mean vs. Texture Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_radius_mean_scatter
```

Radius Mean vs. Texture Mean

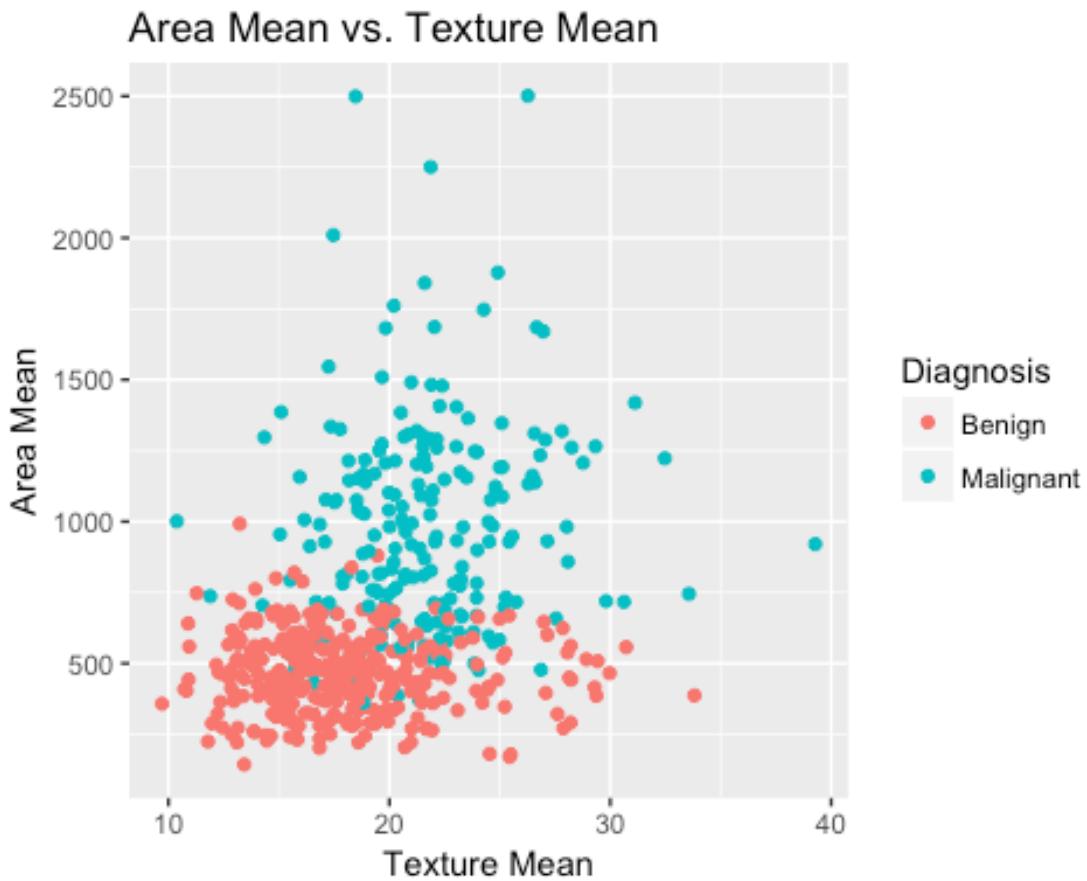


```
#perimeter_mean vs. texture_mean
texture_perimeter_mean_scatter <- ggplot(data = cancer, aes(x = texture_mean,
y = perimeter_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Mean", y = "Perimeter Mean", title = "Perimeter Mean vs. Texture Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_perimeter_mean_scatter
```

Perimeter Mean vs. Texture Mean

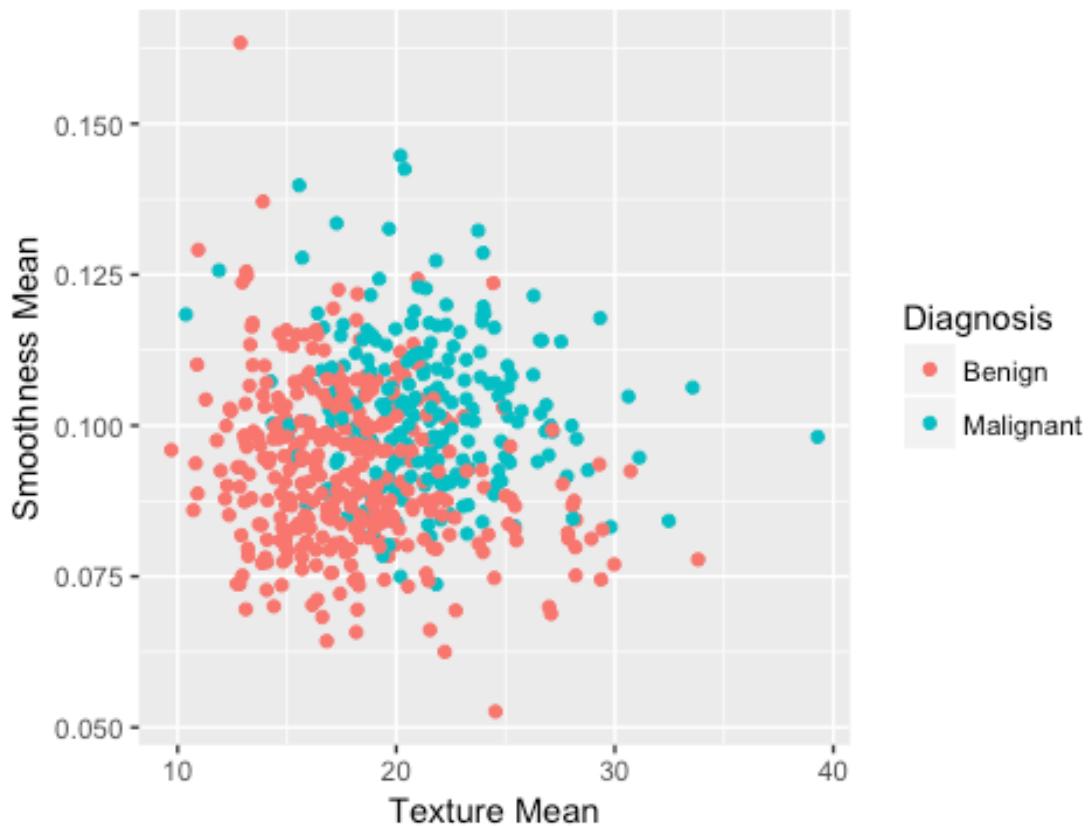


```
#area_mean vs. texture_mean
texture_area_mean_scatter <- ggplot(data = cancer, aes(x = texture_mean, y = area_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Mean", y = "Area Mean", title = "Area Mean vs. Texture Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_area_mean_scatter
```



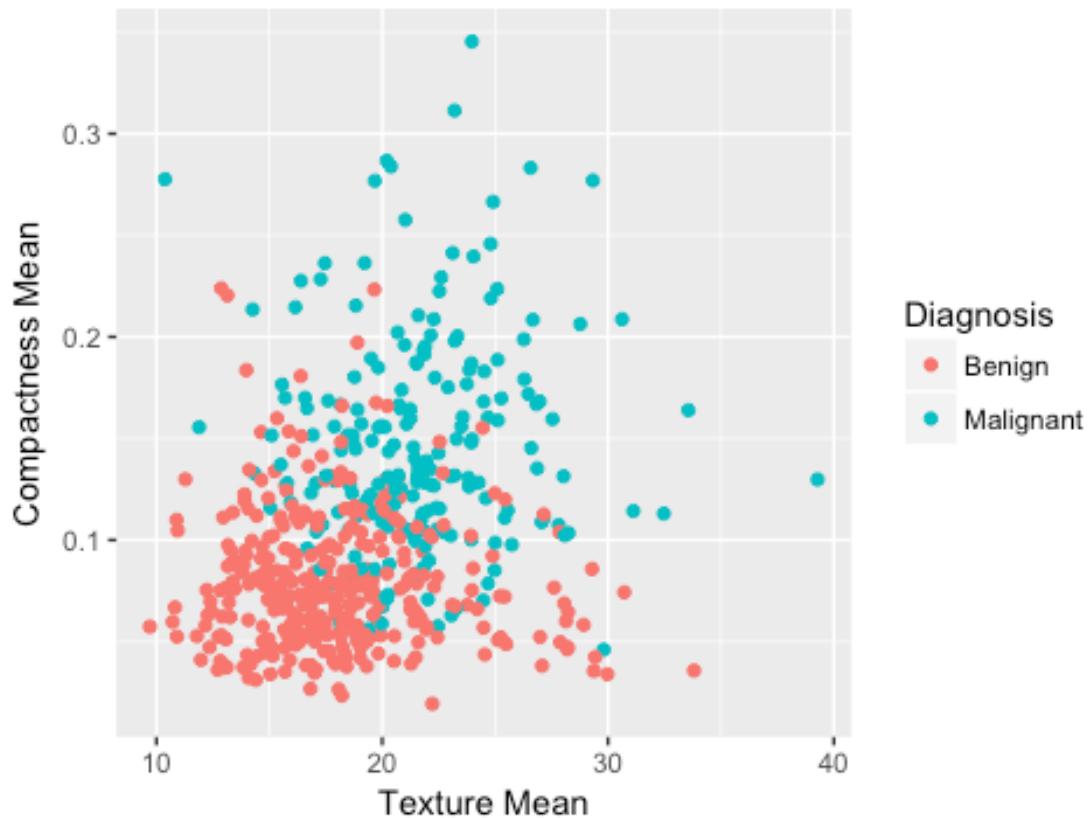
```
#smoothness_mean vs. texture_mean
texture_smoothness_mean_scatter <- ggplot(data = cancer, aes(x = texture_mean,
, y = smoothness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Mean", y = "Smoothness Mean", title = "Smoothness Mean vs. Texture Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_smoothness_mean_scatter
```

Smoothness Mean vs. Texture Mean



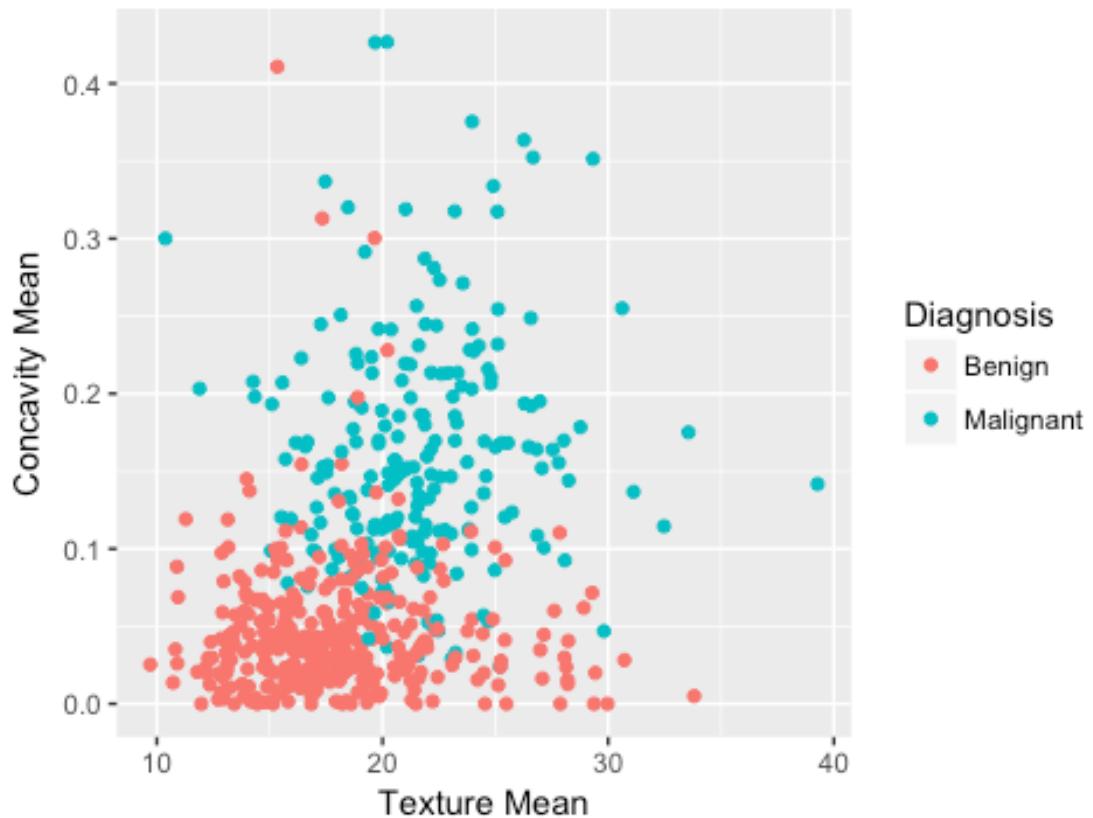
```
#compactness_mean vs. texture_mean
texture_compactness_mean_scatter <- ggplot(data = cancer, aes(x = texture_mean, y = compactness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Mean", y = "Compactness Mean", title = "Compactness Mean vs. Texture Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_compactness_mean_scatter
```

Compactness Mean vs. Texture Mean



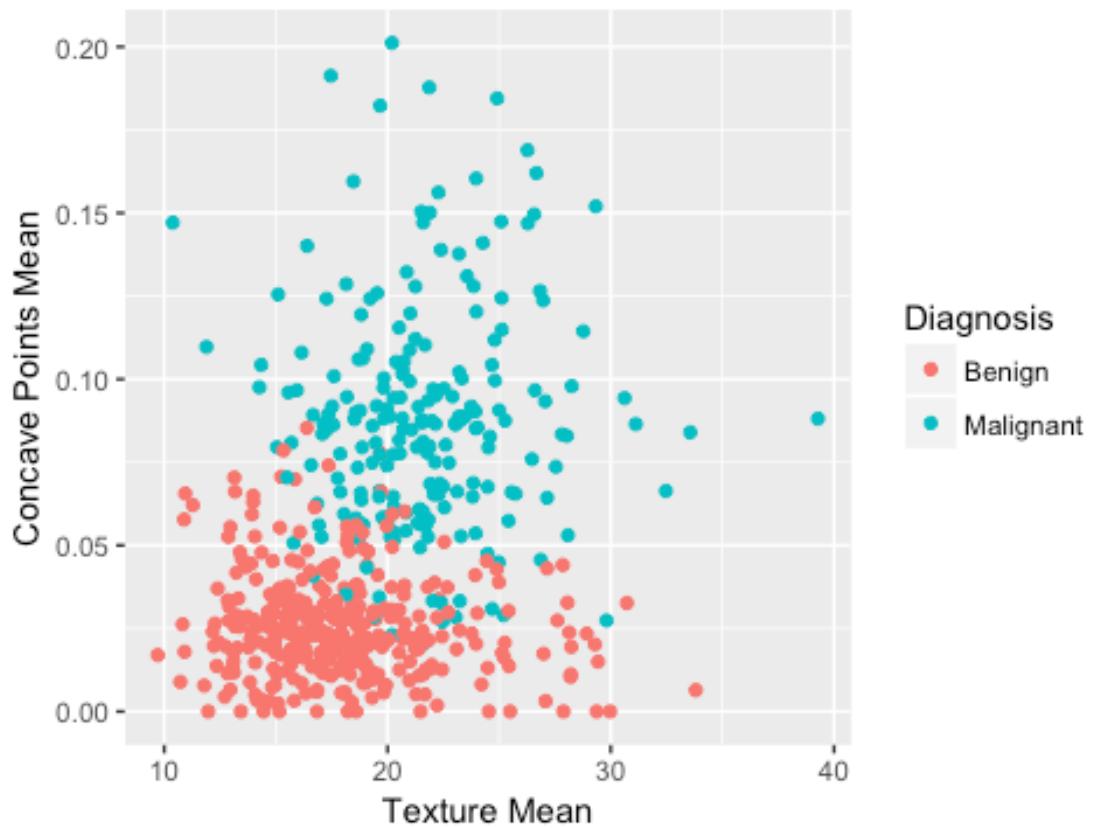
```
#concavity_mean vs. texture_mean
texture_concavity_mean_scatter <- ggplot(data = cancer, aes(x = texture_mean,
y = concavity_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Mean", y = "Concavity Mean", title = "Concavity Mean vs. Texture Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_concavity_mean_scatter
```

Concavity Mean vs. Texture Mean



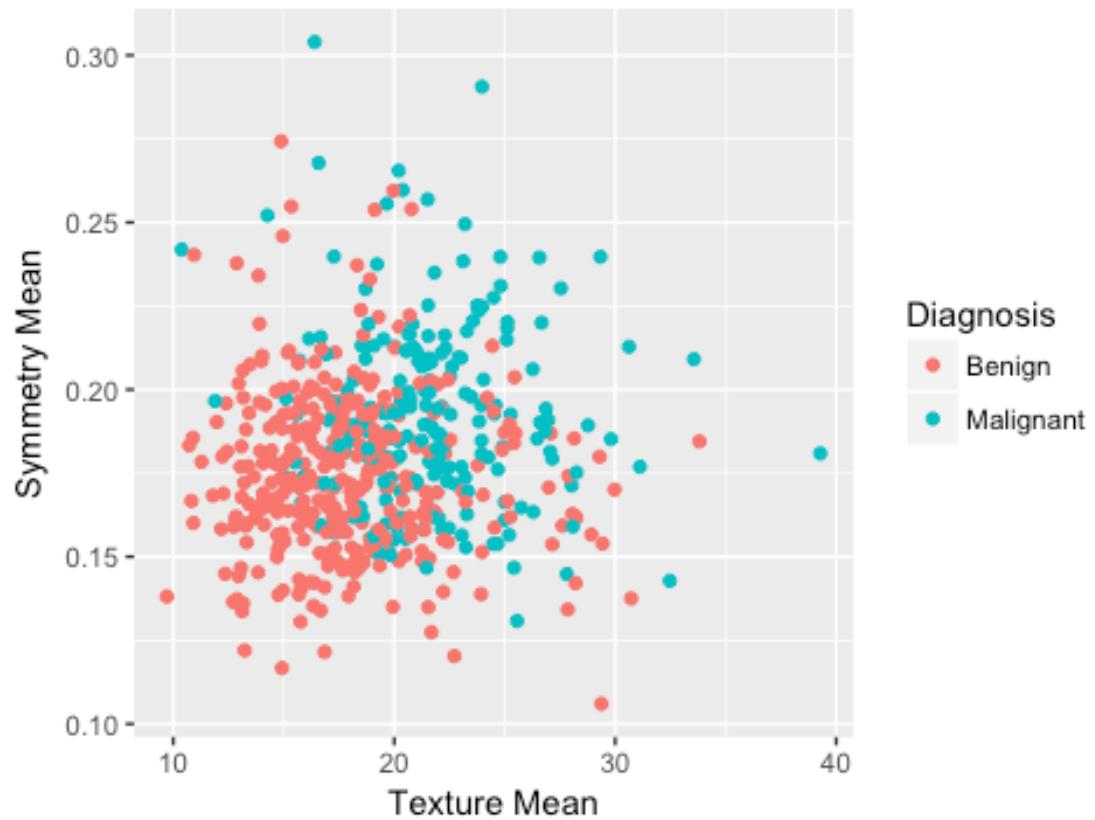
```
#concave.points_mean vs. texture_mean
texture_concave.points_mean_scatter <- ggplot(data = cancer, aes(x = texture_mean, y = concave.points_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Mean", y = "Concave Points Mean", title =
  "Concave Points Mean vs. Texture Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_concave.points_mean_scatter
```

Concave Points Mean vs. Texture Mean



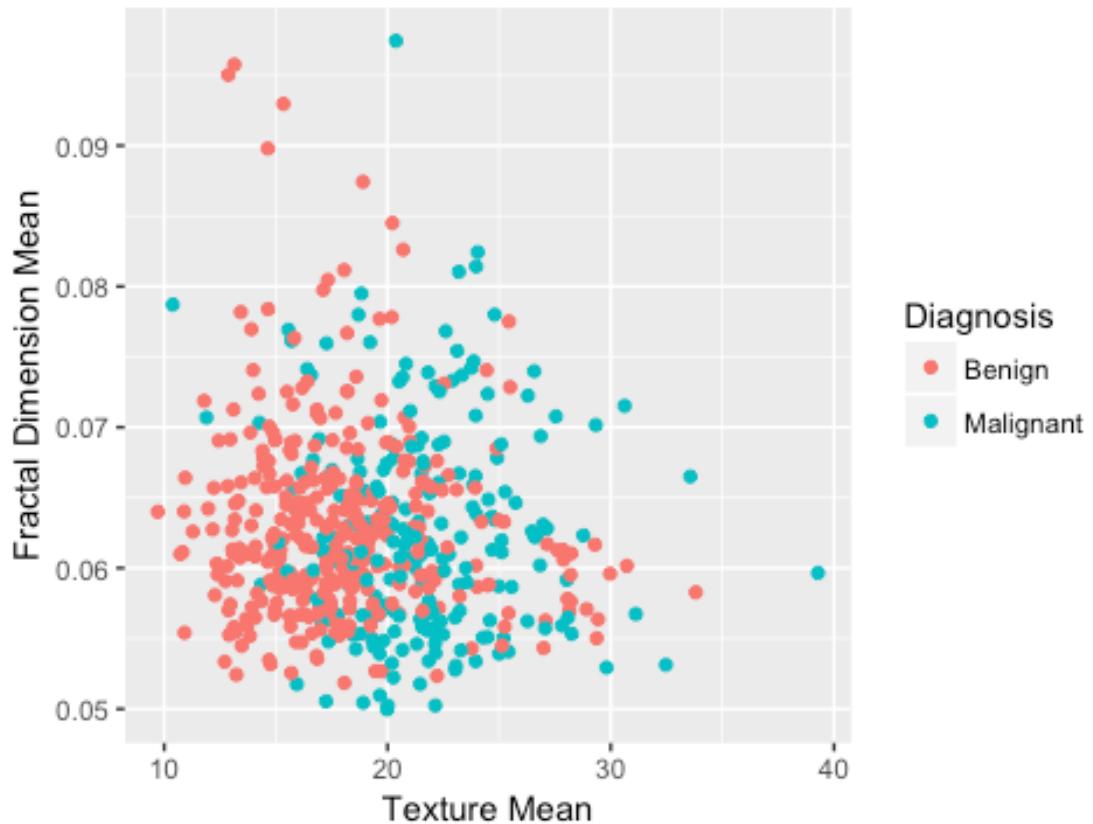
```
#symmetry_mean vs. texture_mean
texture_symmetry_mean_scatter <- ggplot(data = cancer, aes(x = texture_mean,
y = symmetry_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Mean", y = "Symmetry Mean", title = "Symmetry Mean vs. Texture Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_symmetry_mean_scatter
```

Symmetry Mean vs. Texture Mean



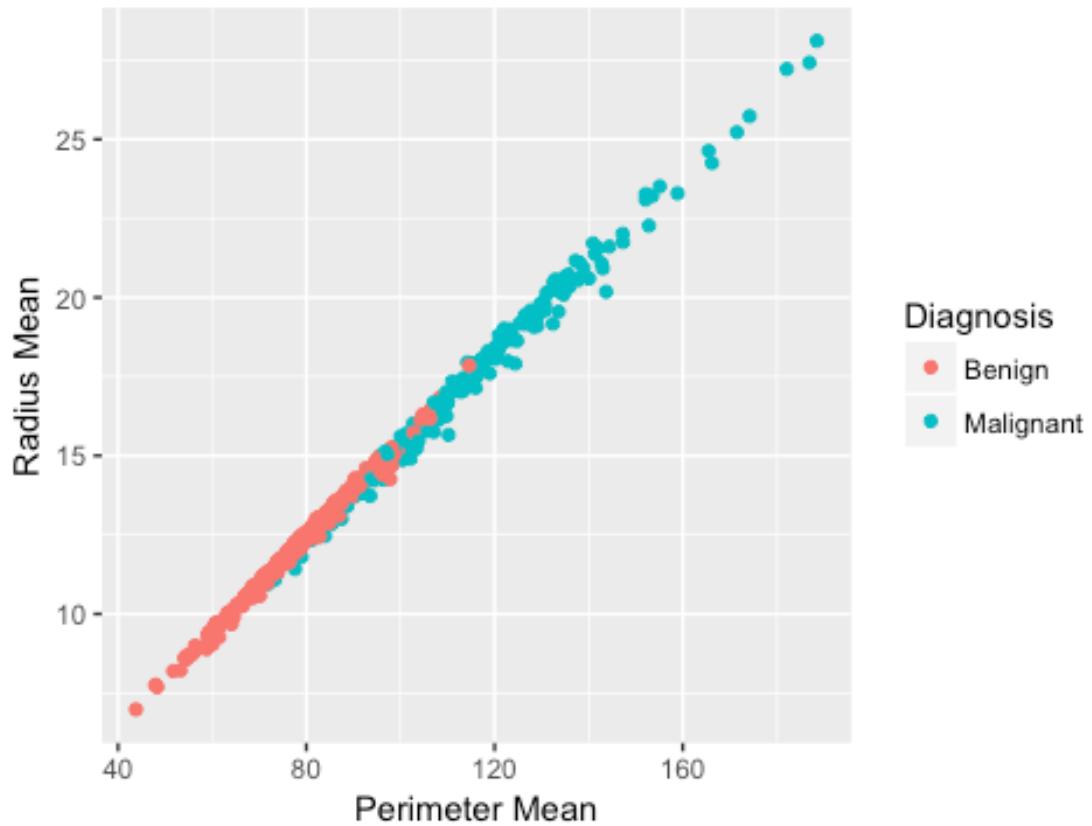
```
#fractal_dimension_mean vs. texture_mean
texture_fractal_dimension_mean_scatter <- ggplot(data = cancer, aes(x = texture_mean, y = fractal_dimension_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Mean", y = "Fractal Dimension Mean", title =
  "Fractal Dimension Mean vs. Texture Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_fractal_dimension_mean_scatter
```

Fractal Dimension Mean vs. Texture Mean



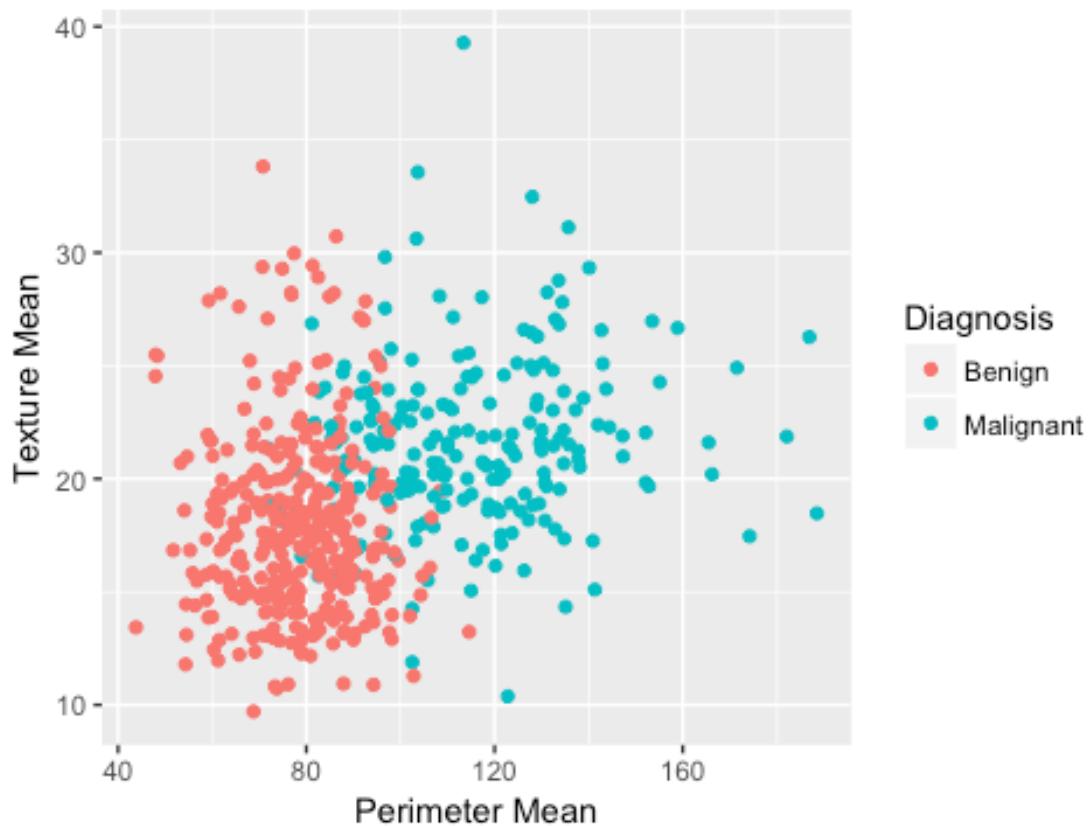
```
#radius_mean vs. perimeter_mean
perimeter_radius_mean_scatter <- ggplot(data = cancer, aes(x = perimeter_mean,
, y = radius_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Mean", y = "Radius Mean", title = "Radius Mean vs. Perimeter Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_radius_mean_scatter
```

Radius Mean vs. Perimeter Mean



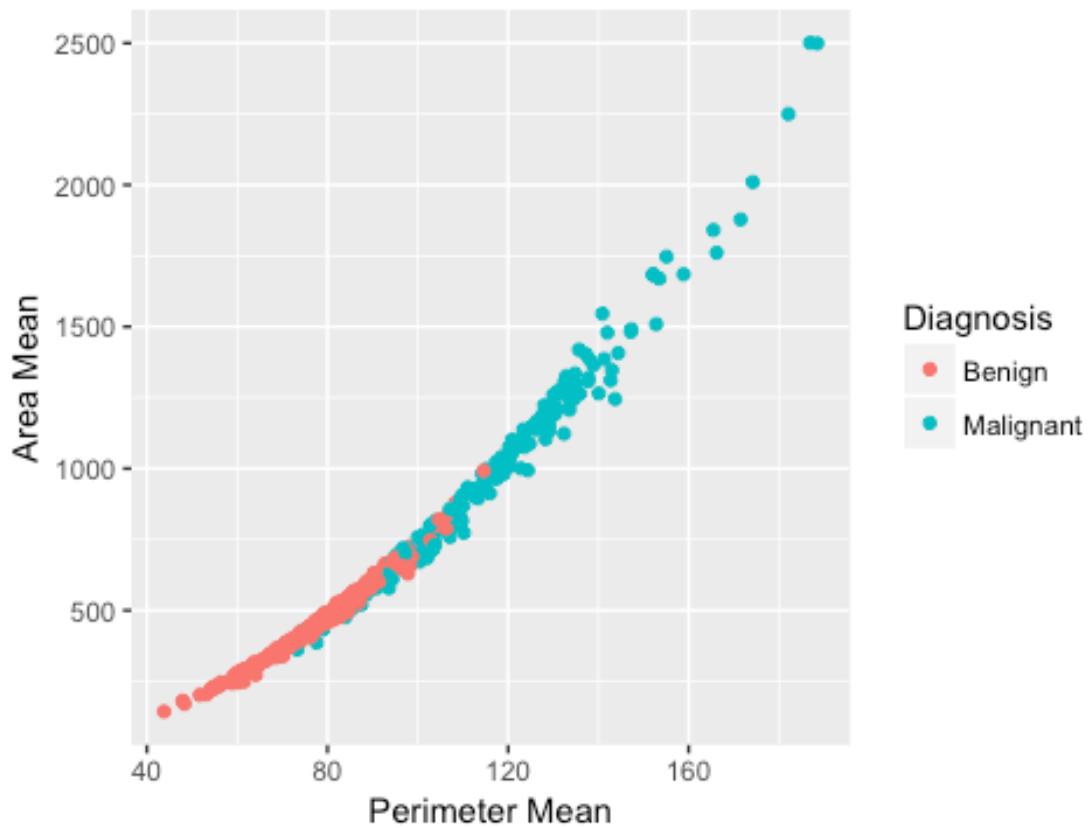
```
#texture_mean vs. perimeter_mean
perimeter_texture_mean_scatter <- ggplot(data = cancer, aes(x = perimeter_mean, y = texture_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Mean", y = "Texture Mean", title = "Texture Mean vs. Perimeter Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_texture_mean_scatter
```

Texture Mean vs. Perimeter Mean



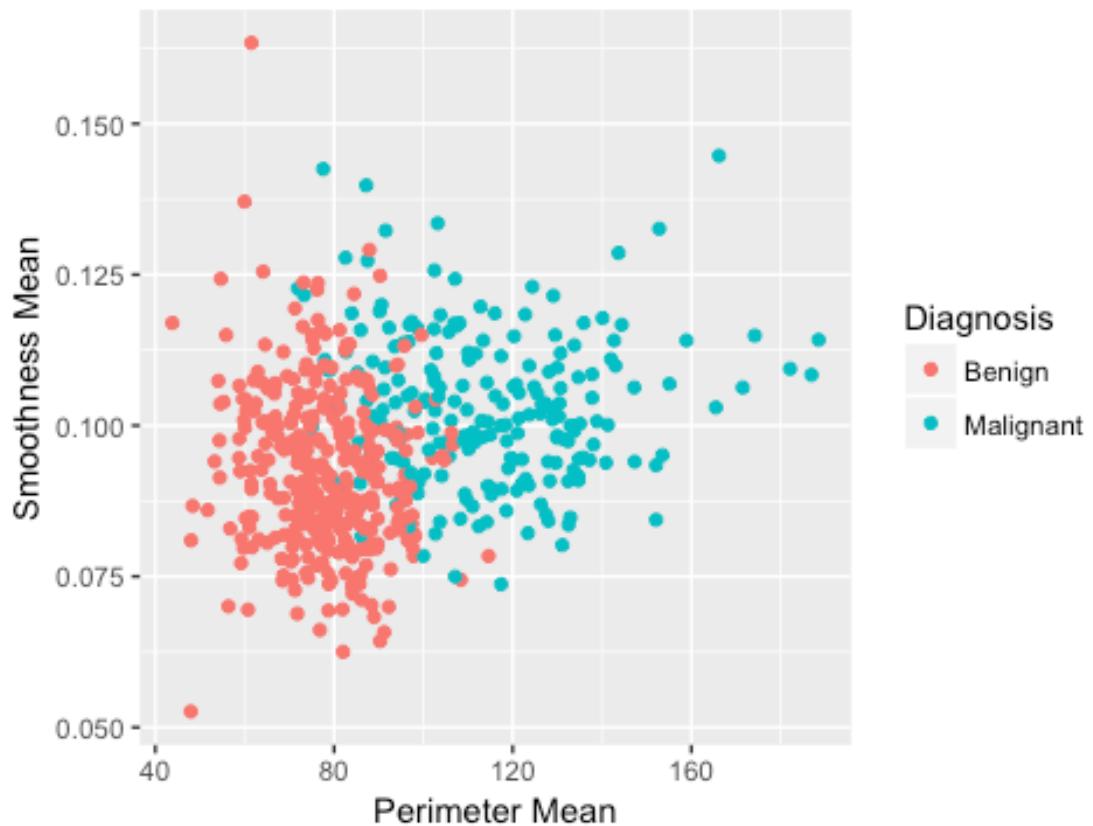
```
#area_mean vs. perimeter_mean
perimeter_area_mean_scatter <- ggplot(data = cancer, aes(x = perimeter_mean,
y = area_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Mean", y = "Area Mean", title = "Area Mean vs. Perimeter Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_area_mean_scatter
```

Area Mean vs. Perimeter Mean



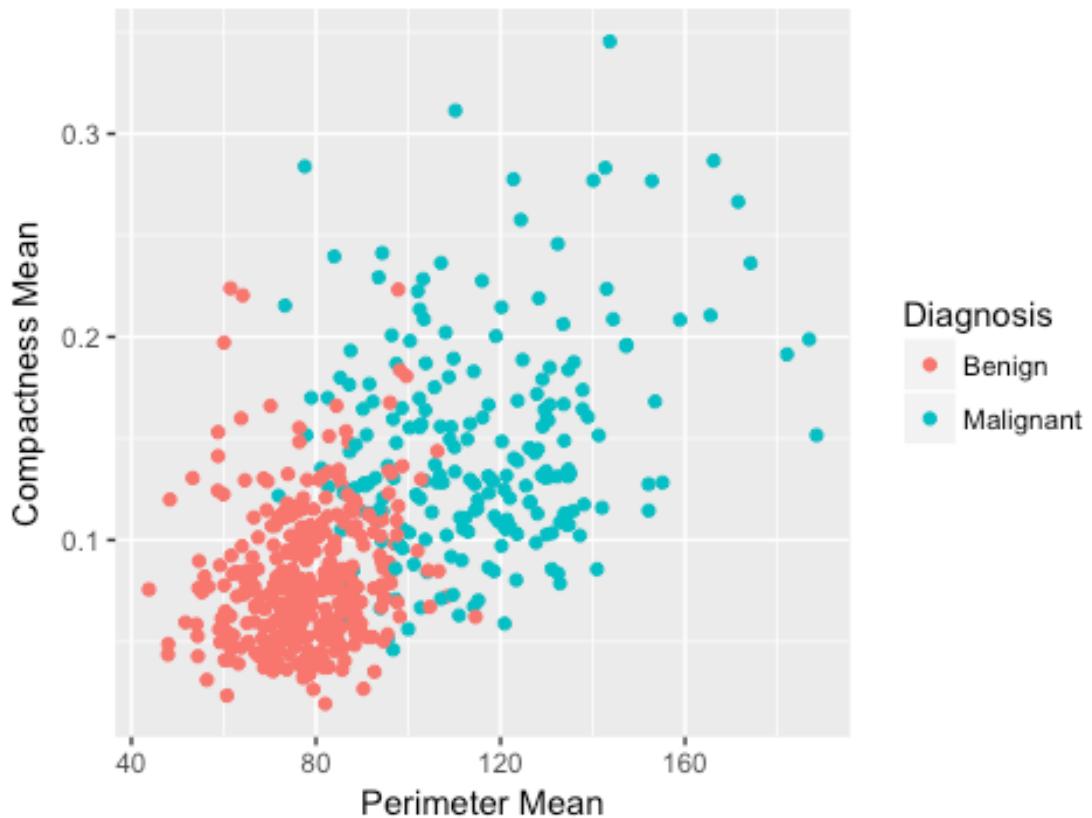
```
#smoothness_mean vs. perimeter_mean
perimeter_smoothness_mean_scatter <- ggplot(data = cancer, aes(x = perimeter_mean, y = smoothness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Mean", y = "Smoothness Mean", title = "Smoothness Mean vs. Perimeter Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_smoothness_mean_scatter
```

Smoothness Mean vs. Perimeter Mean



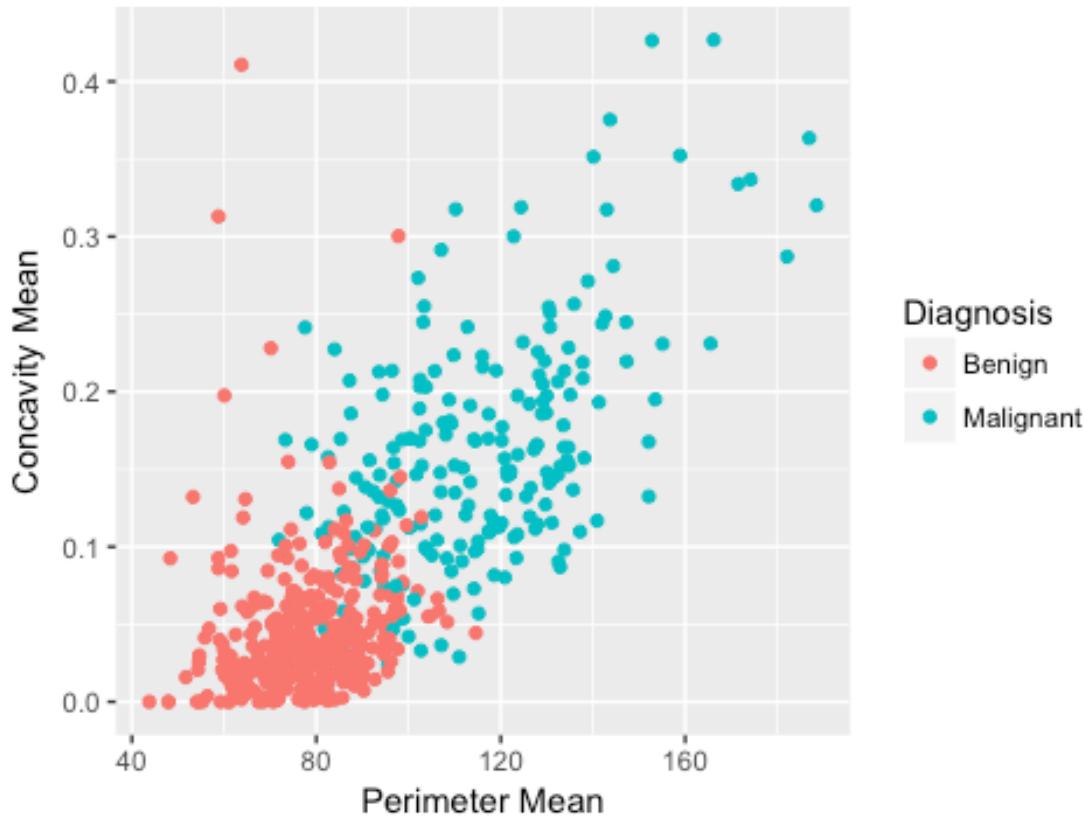
```
#compactness_mean vs. perimeter_mean
perimeter_compactness_mean_scatter <- ggplot(data = cancer, aes(x = perimeter_mean, y = compactness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Mean", y = "Compactness Mean", title = "Compactness Mean vs. Perimeter Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_compactness_mean_scatter
```

Compactness Mean vs. Perimeter Mean



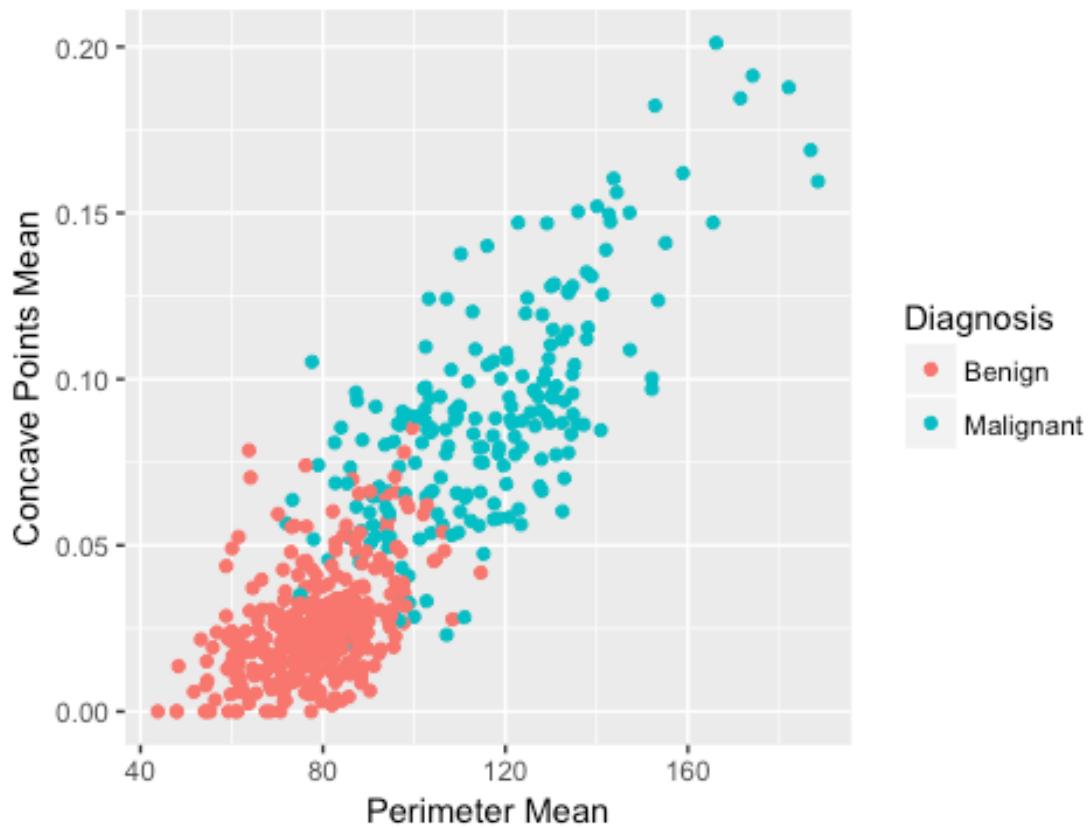
```
#concavity_mean vs. perimeter_mean
perimeter_concavity_mean_scatter <- ggplot(data = cancer, aes(x = perimeter_mean, y = concavity_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Mean", y = "Concavity Mean", title = "Concavity Mean vs. Perimeter Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_concavity_mean_scatter
```

Concavity Mean vs. Perimeter Mean



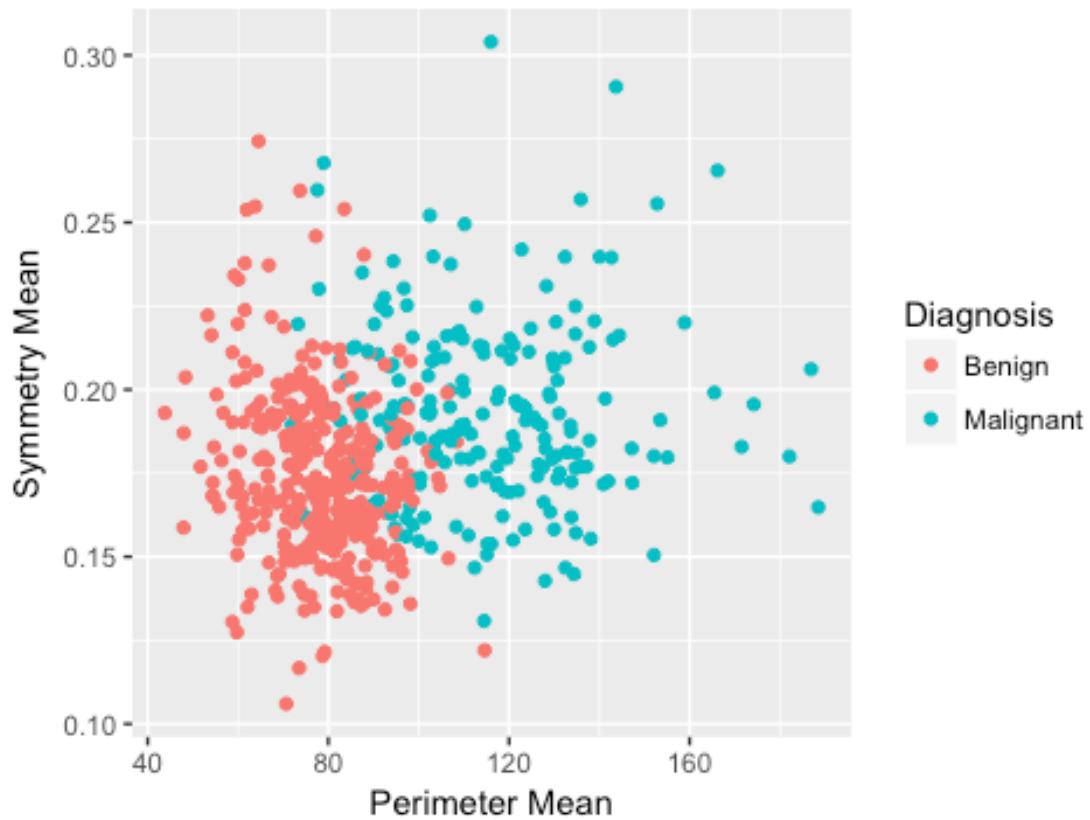
```
#concave.points_mean vs. perimeter_mean
perimeter_concave.points_mean_scatter <- ggplot(data = cancer, aes(x = perimeter_mean, y = concave.points_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Mean", y = "Concave Points Mean", title =
  "Concave Points Mean vs. Perimeter Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_concave.points_mean_scatter
```

Concave Points Mean vs. Perimeter Mean



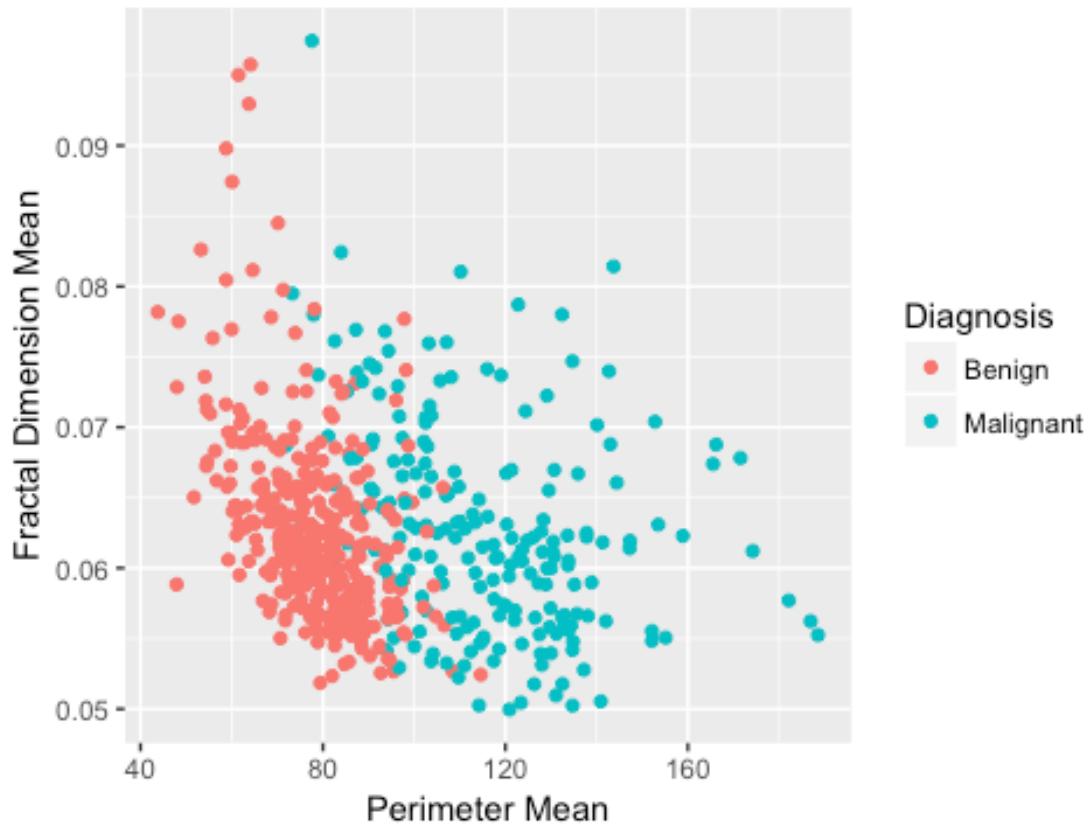
```
#symmetry_mean vs. perimeter_mean
perimeter_symmetry_mean_scatter <- ggplot(data = cancer, aes(x = perimeter_mean, y = symmetry_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Mean", y = "Symmetry Mean", title = "Symmetry Mean vs. Perimeter Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_symmetry_mean_scatter
```

Symmetry Mean vs. Perimeter Mean



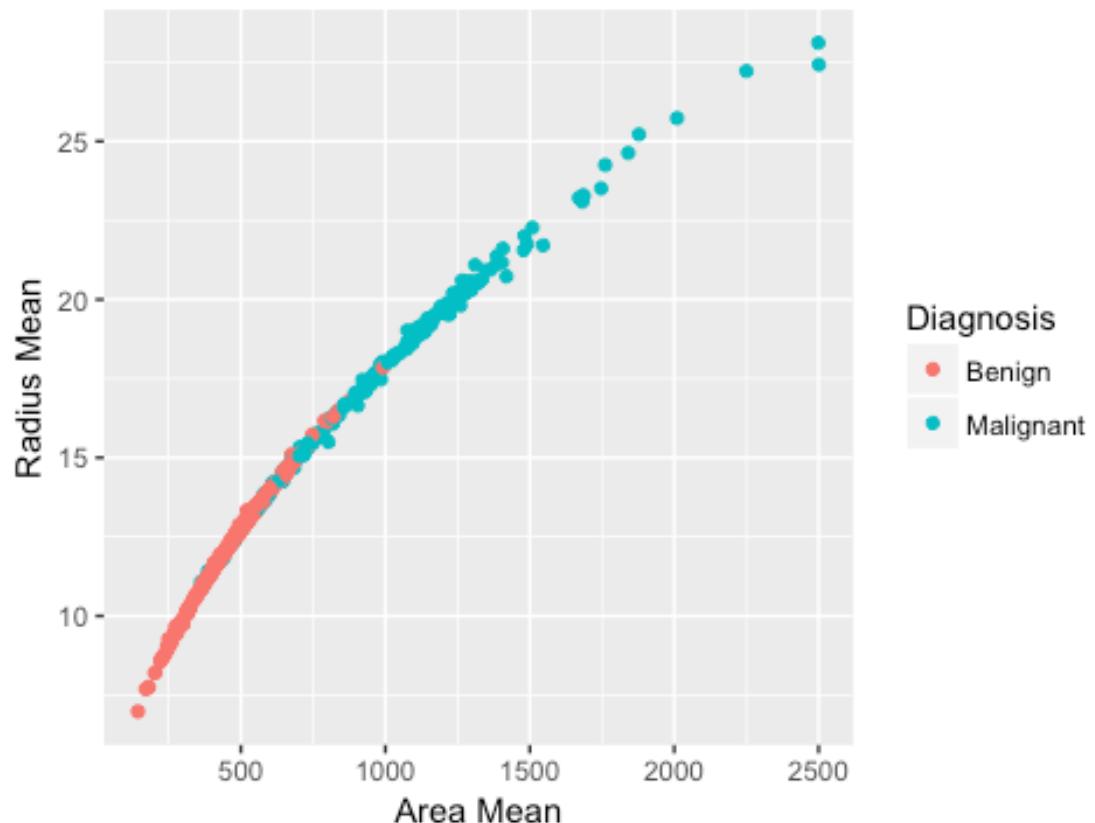
```
#fractal_dimension_mean vs. perimeter_mean
perimeter_fractal_dimension_mean_scatter <- ggplot(data = cancer, aes(x = perimeter_mean, y = fractal_dimension_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Mean", y = "Fractal Dimension Mean", title = "Fractal Dimension Mean vs. Perimeter Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_fractal_dimension_mean_scatter
```

Fractal Dimension Mean vs. Perimeter Mean



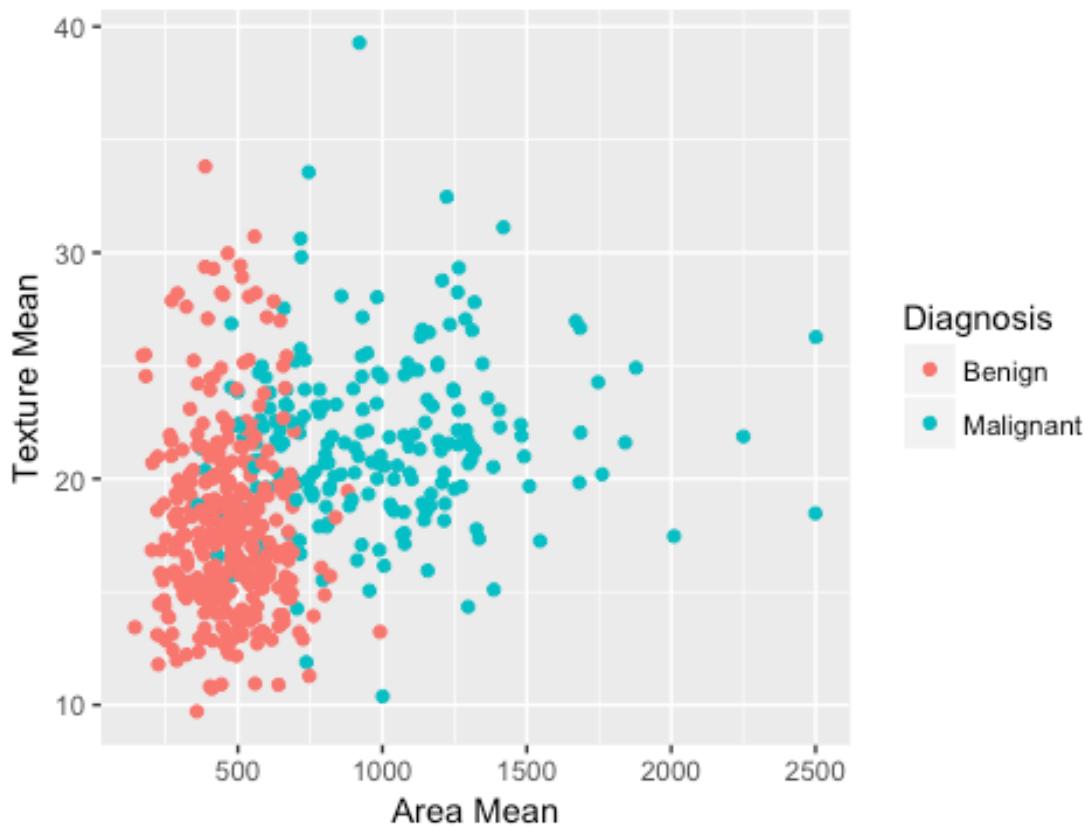
```
#radius_mean vs. area_mean
area_radius_mean_scatter <- ggplot(data = cancer, aes(x = area_mean, y = radius_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Mean", y = "Radius Mean", title = "Radius Mean vs. Area Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_radius_mean_scatter
```

Radius Mean vs. Area Mean



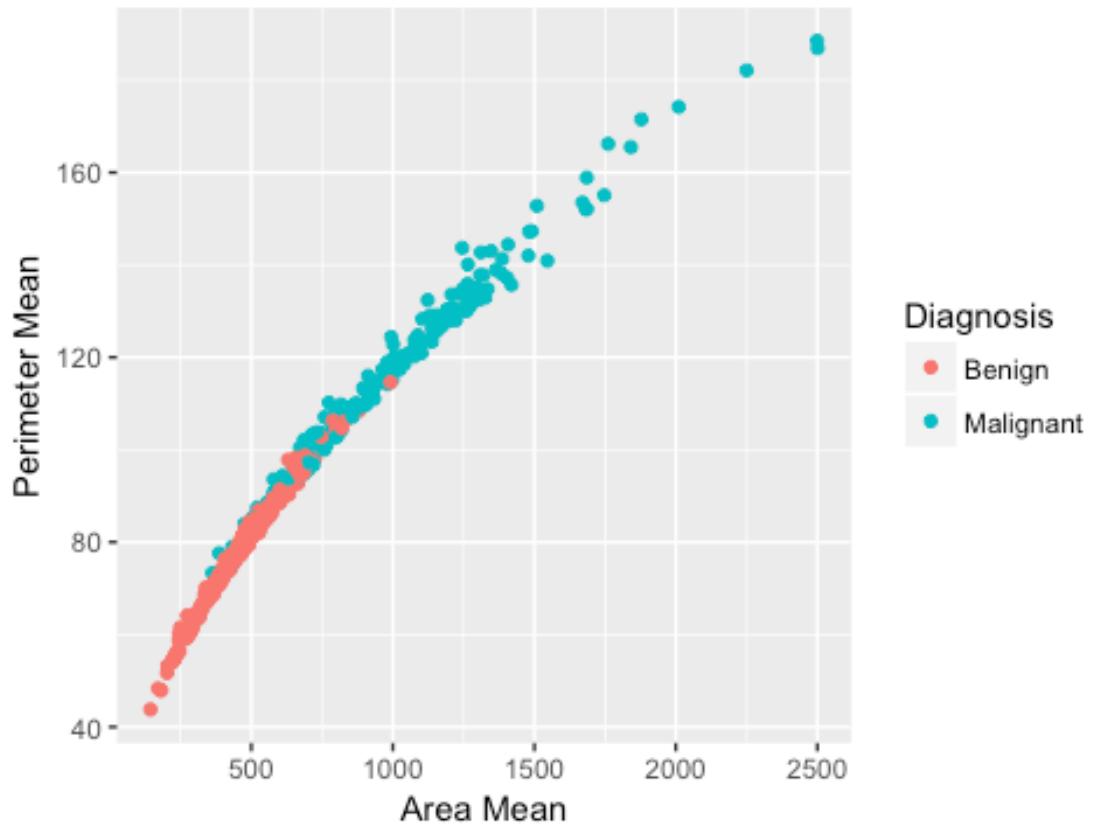
```
#texture_mean vs. area_mean
area_texture_mean_scatter <- ggplot(data = cancer, aes(x = area_mean, y = texture_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Mean", y = "Texture Mean", title = "Texture Mean vs. Area Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_texture_mean_scatter
```

Texture Mean vs. Area Mean



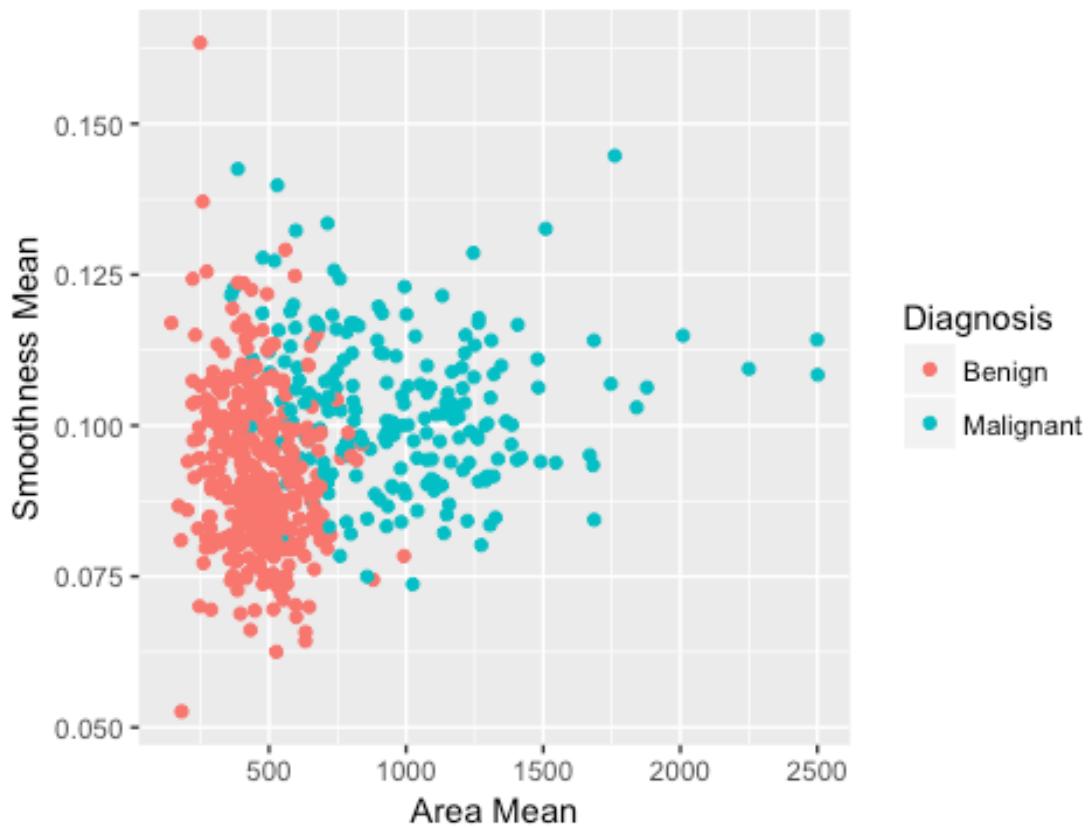
```
#perimeter_mean vs. area_mean
area_perimeter_mean_scatter <- ggplot(data = cancer, aes(x = area_mean, y = perimeter_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Mean", y = "Perimeter Mean", title = "Perimeter Mean vs. Area Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_perimeter_mean_scatter
```

Perimeter Mean vs. Area Mean



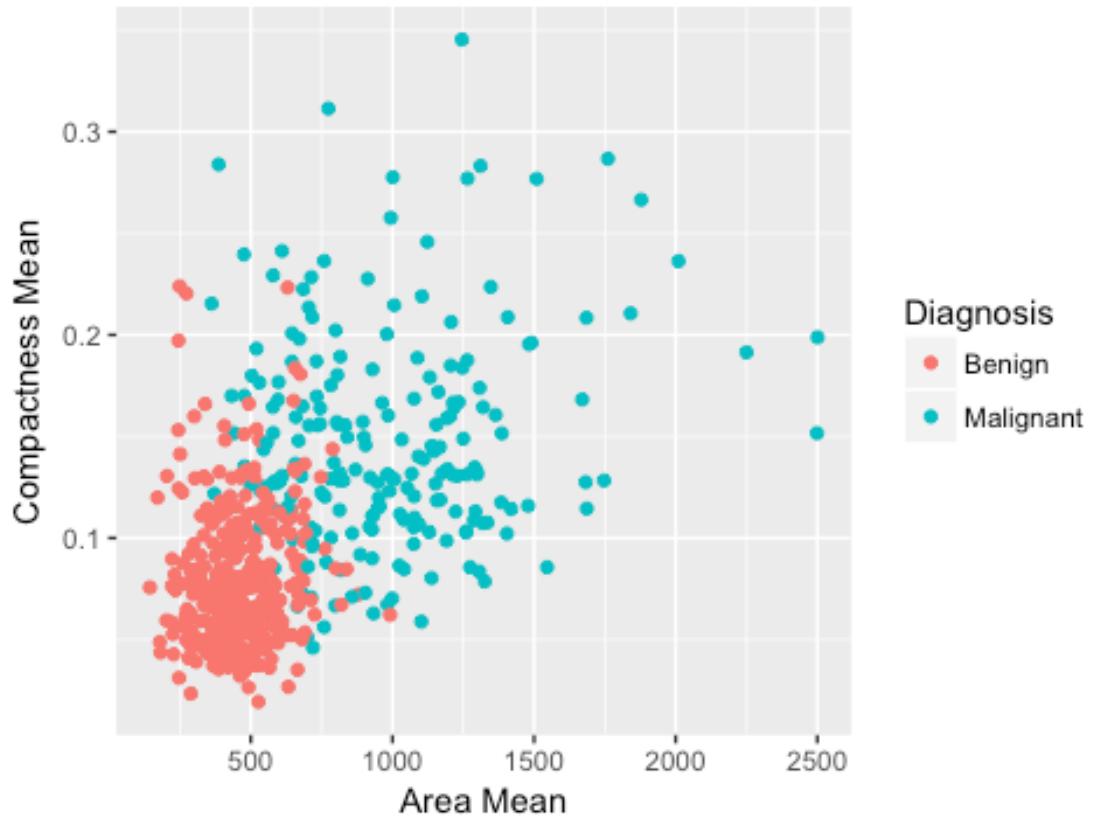
```
#smoothness_mean vs. area_mean
area_smoothness_mean_scatter <- ggplot(data = cancer, aes(x = area_mean, y = smoothness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Mean", y = "Smoothness Mean", title = "Smoothness Mean vs. Area Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_smoothness_mean_scatter
```

Smoothness Mean vs. Area Mean



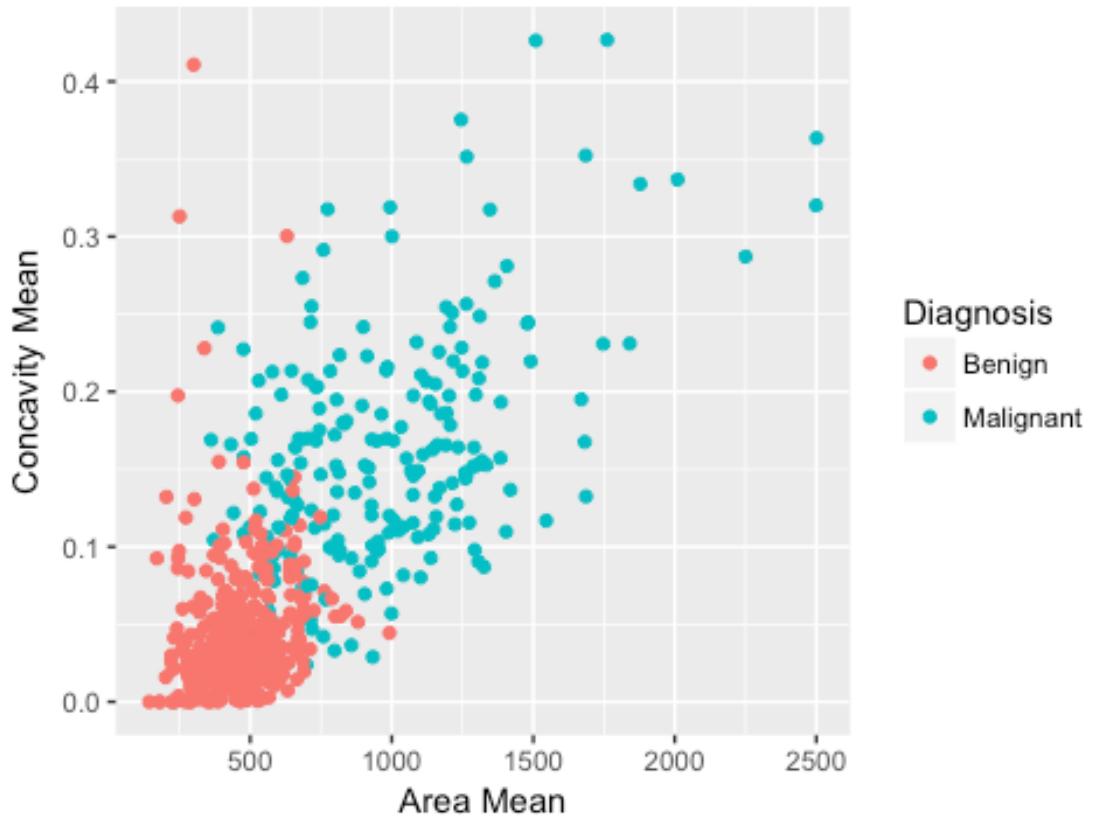
```
#compactness_mean vs. area_mean
area_compactness_mean_scatter <- ggplot(data = cancer, aes(x = area_mean, y = compactness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Mean", y = "Compactness Mean", title = "Compactness Mean vs. Area Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_compactness_mean_scatter
```

Compactness Mean vs. Area Mean

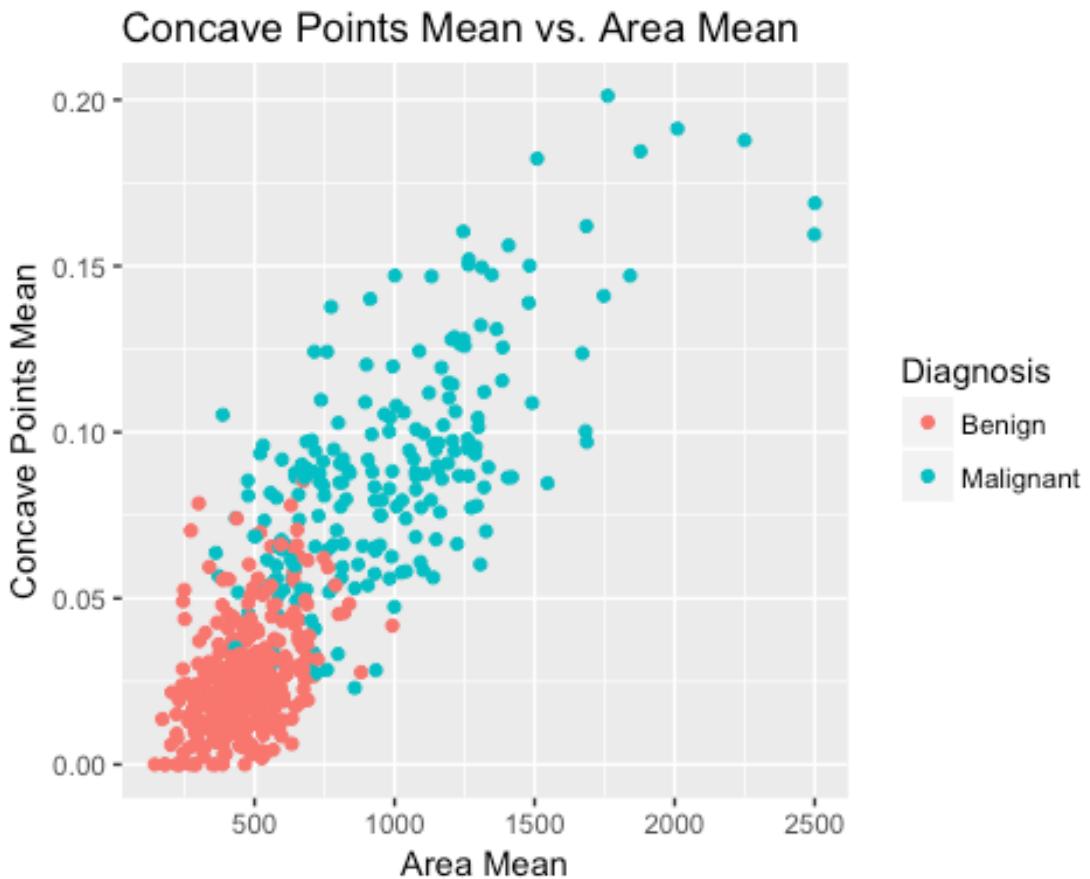


```
#concavity_mean vs. area_mean
area_concavity_mean_scatter <- ggplot(data = cancer, aes(x = area_mean, y = concavity_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Mean", y = "Concavity Mean", title = "Concavity Mean vs. Area Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_concavity_mean_scatter
```

Concavity Mean vs. Area Mean

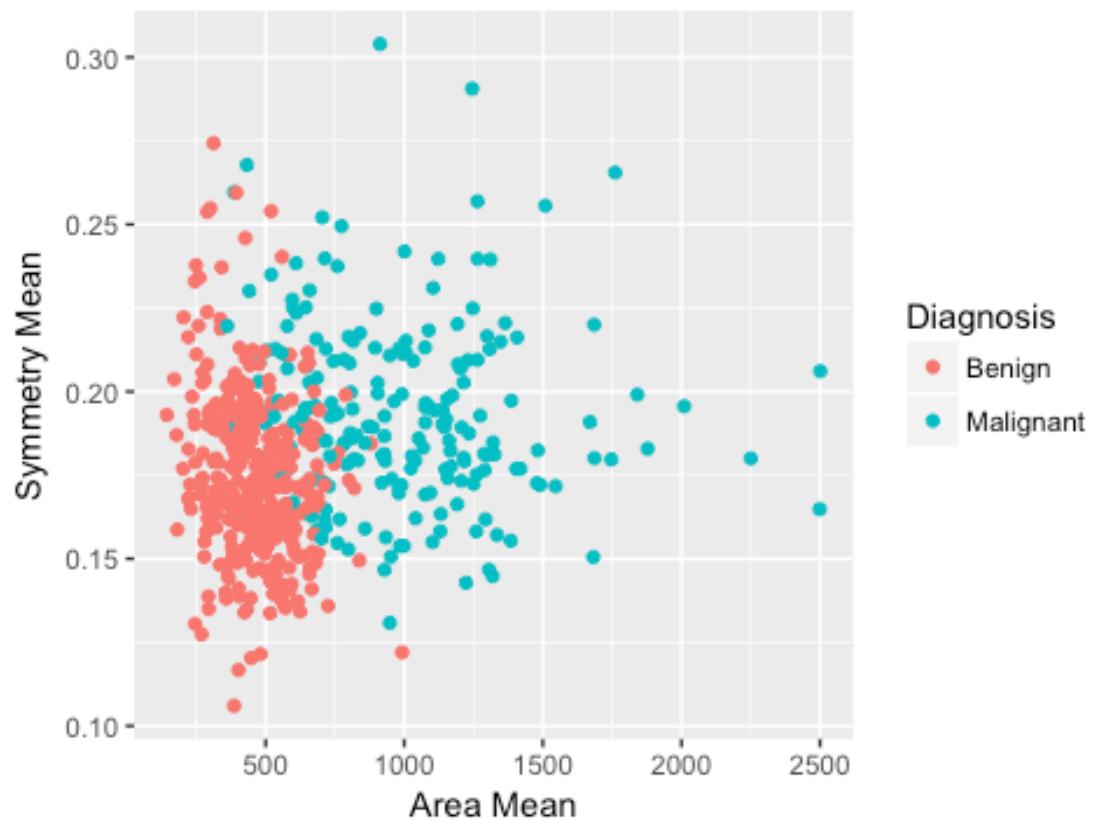


```
#concave.points_mean vs. area_mean
area_concave.points_mean_scatter <- ggplot(data = cancer, aes(x = area_mean,
y = concave.points_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Mean", y = "Concave Points Mean", title = "Concave Points Mean vs. Area Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_concave.points_mean_scatter
```



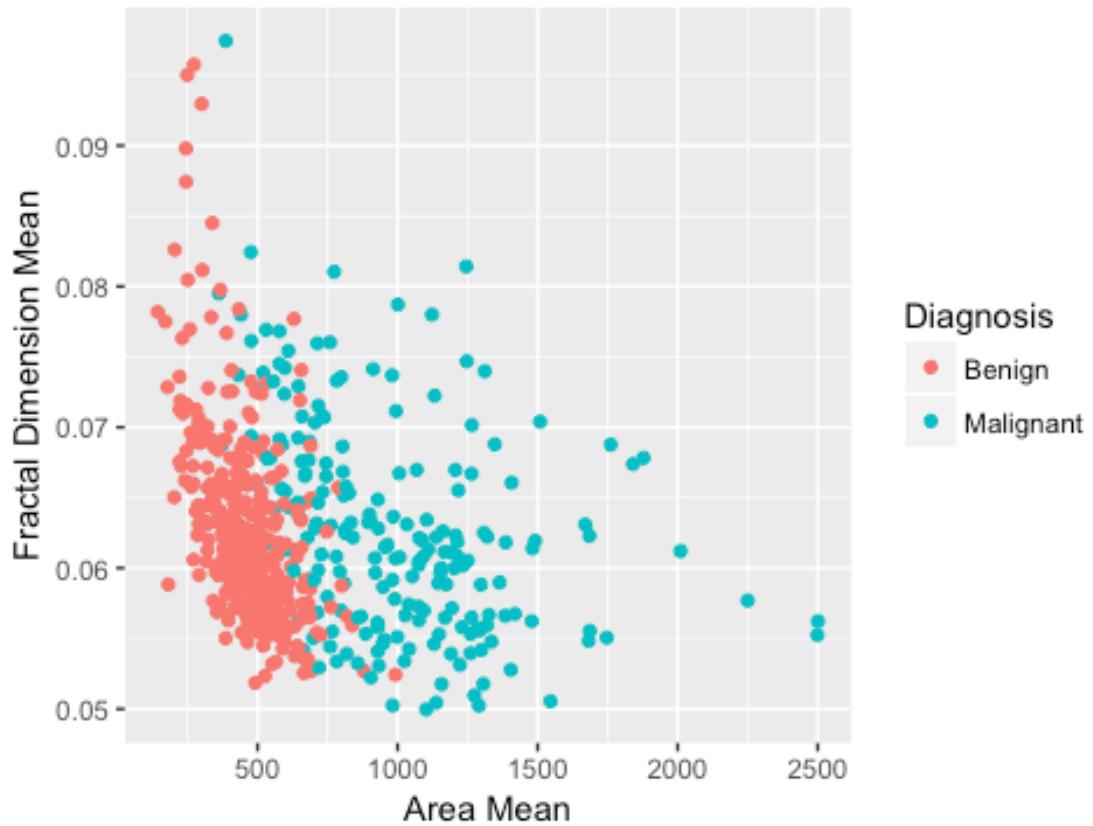
```
#symmetry_mean vs. area_mean
area_symmetry_mean_scatter <- ggplot(data = cancer, aes(x = area_mean, y = symmetry_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Mean", y = "Symmetry Mean", title = "Symmetry Mean vs. Area Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_symmetry_mean_scatter
```

Symmetry Mean vs. Area Mean



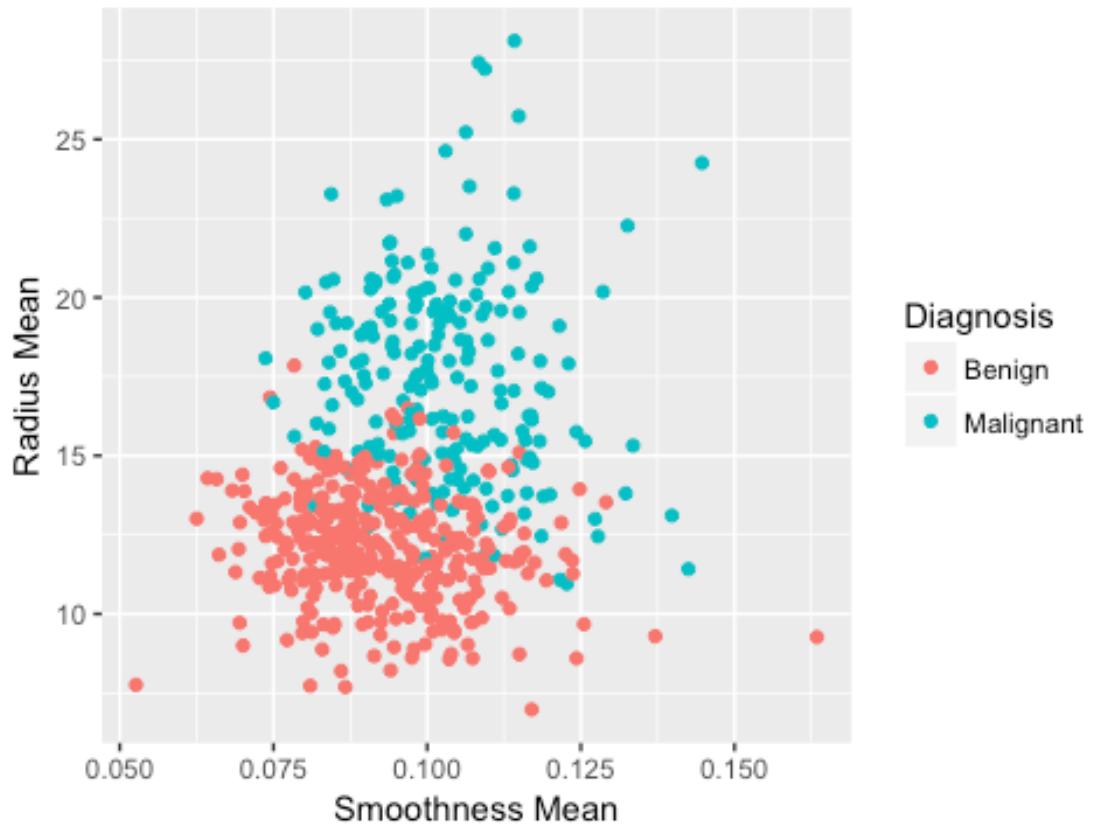
```
#fractal_dimension_mean vs. area_mean
area_fractal_dimension_mean_scatter <- ggplot(data = cancer, aes(x = area_mean, y = fractal_dimension_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Mean", y = "Fractal Dimension Mean", title =
  "Fractal Dimension Mean vs. Area Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_fractal_dimension_mean_scatter
```

Fractal Dimension Mean vs. Area Mean



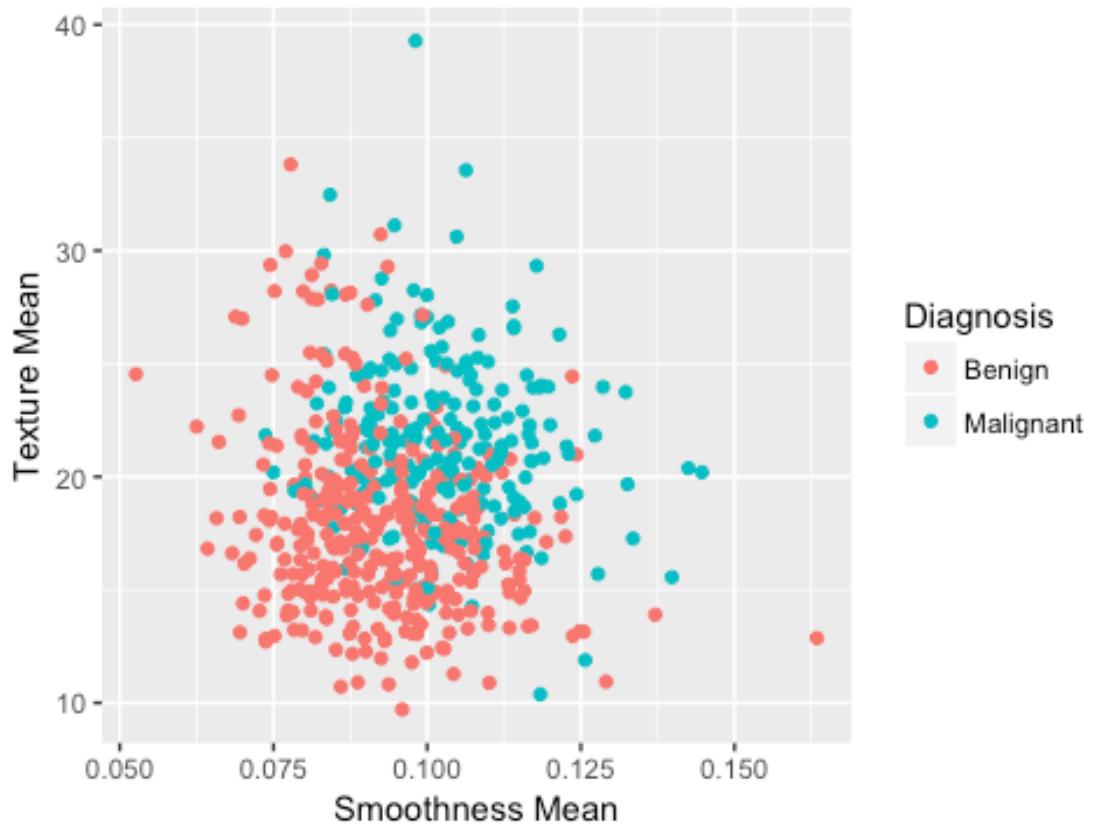
```
#radius_mean vs. smoothness_mean
smoothness_radius_mean_scatter <- ggplot(data = cancer, aes(x = smoothness_mean, y = radius_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Mean", y = "Radius Mean", title = "Radius Mean vs. Smoothness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_radius_mean_scatter
```

Radius Mean vs. Smoothness Mean



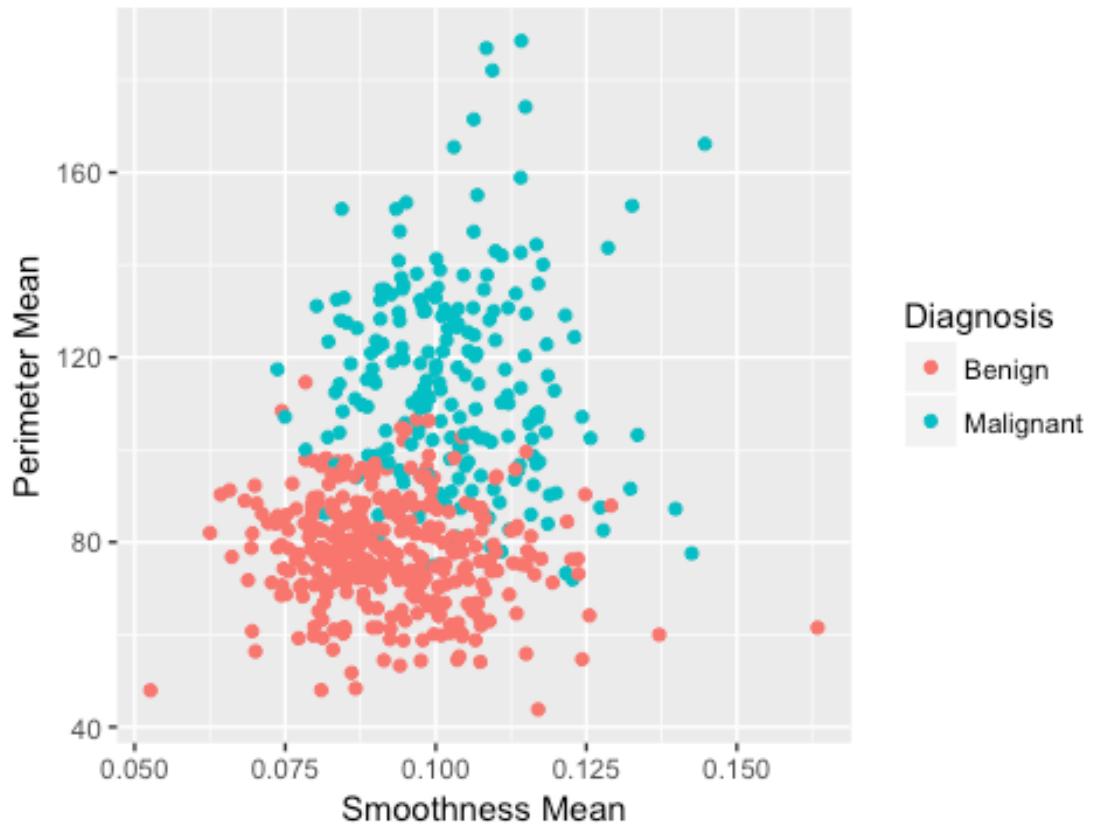
```
#texture_mean vs. smoothness_mean
smoothness_texture_mean_scatter <- ggplot(data = cancer, aes(x = smoothness_mean, y = texture_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Mean", y = "Texture Mean", title = "Texture Mean vs. Smoothness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_texture_mean_scatter
```

Texture Mean vs. Smoothness Mean



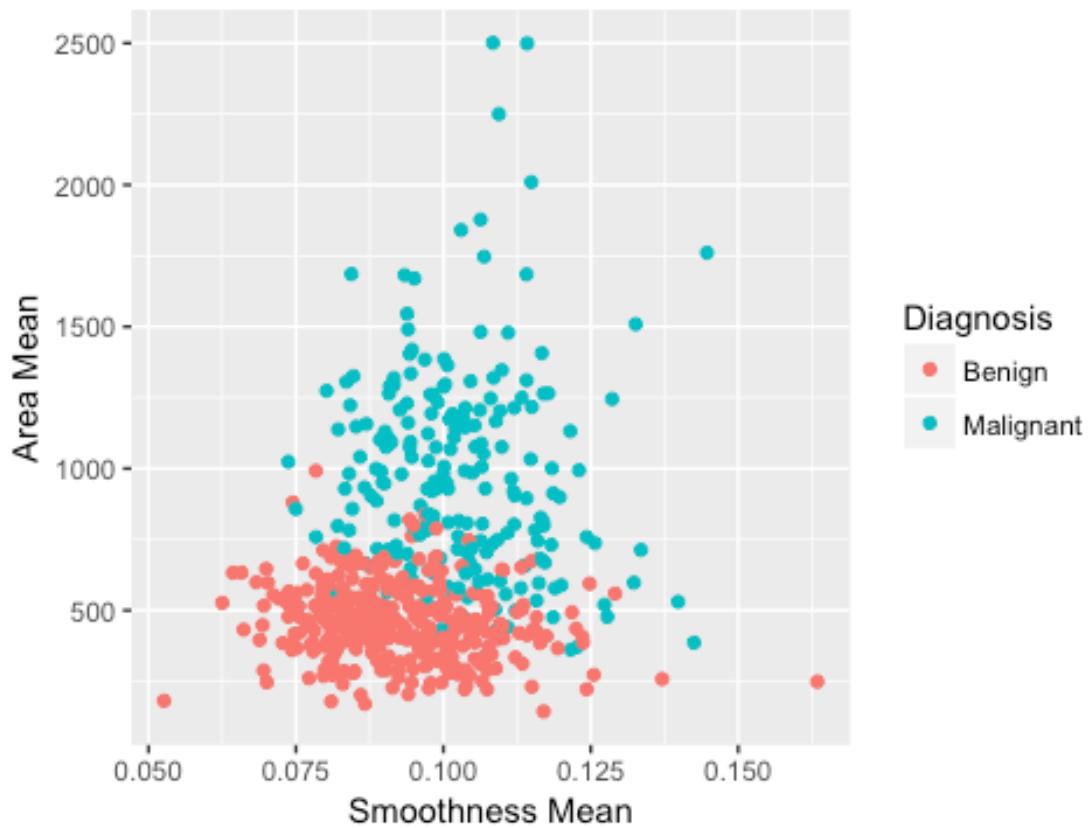
```
#perimeter_mean vs. smoothness_mean
smoothness_perimeter_mean_scatter <- ggplot(data = cancer, aes(x = smoothness_mean, y = perimeter_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Mean", y = "Perimeter Mean", title = "Perimeter Mean vs. Smoothness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_perimeter_mean_scatter
```

Perimeter Mean vs. Smoothness Mean



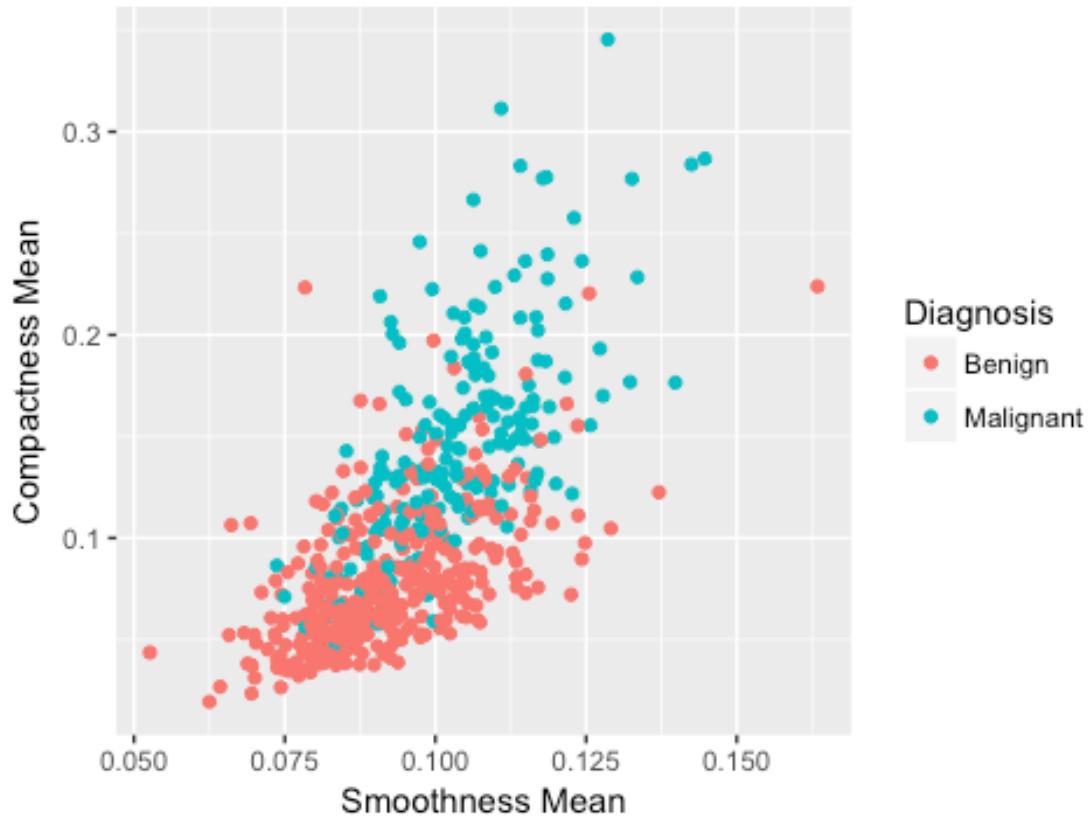
```
#area_mean vs. smoothness_mean
smoothness_area_mean_scatter <- ggplot(data = cancer, aes(x = smoothness_mean,
, y = area_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Mean", y = "Area Mean", title = "Area Mean vs. Smoothness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_area_mean_scatter
```

Area Mean vs. Smoothness Mean



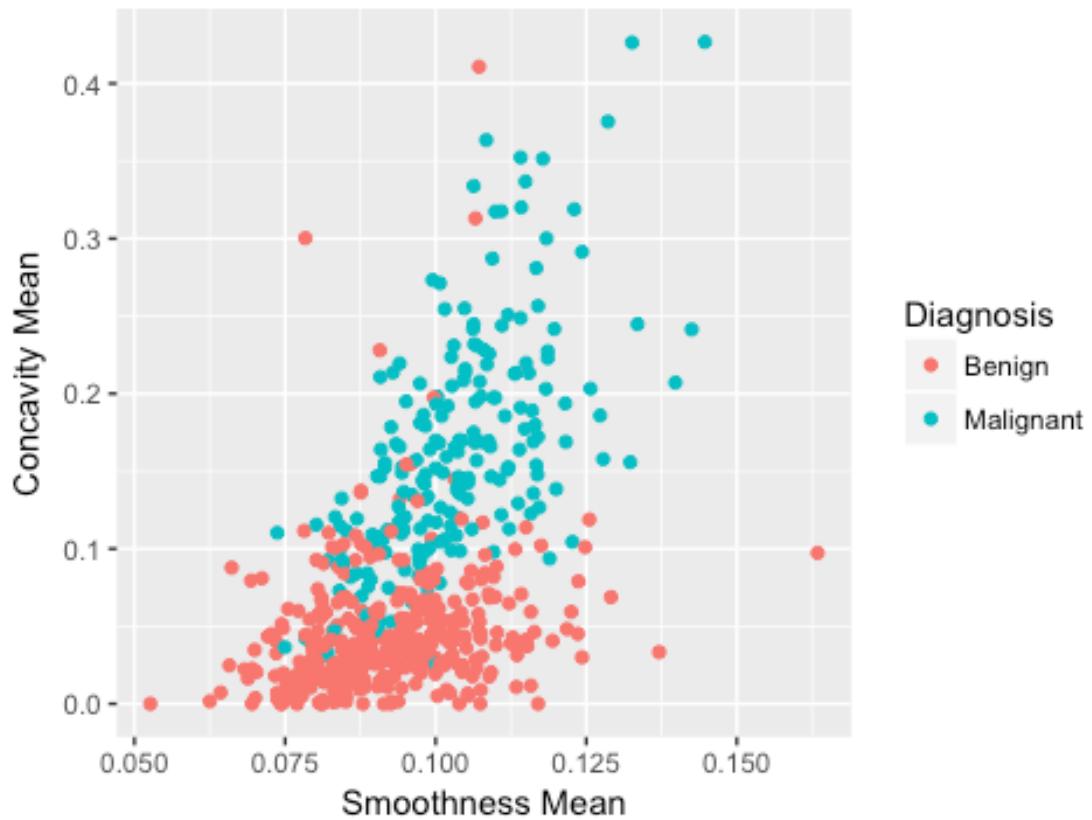
```
#compactness_mean vs. smoothness_mean
smoothness_compactness_mean_scatter <- ggplot(data = cancer, aes(x = smoothness_mean, y = compactness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Mean", y = "Compactness Mean", title =
  "Compactness Mean vs. Smoothness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_compactness_mean_scatter
```

Compactness Mean vs. Smoothness Mean

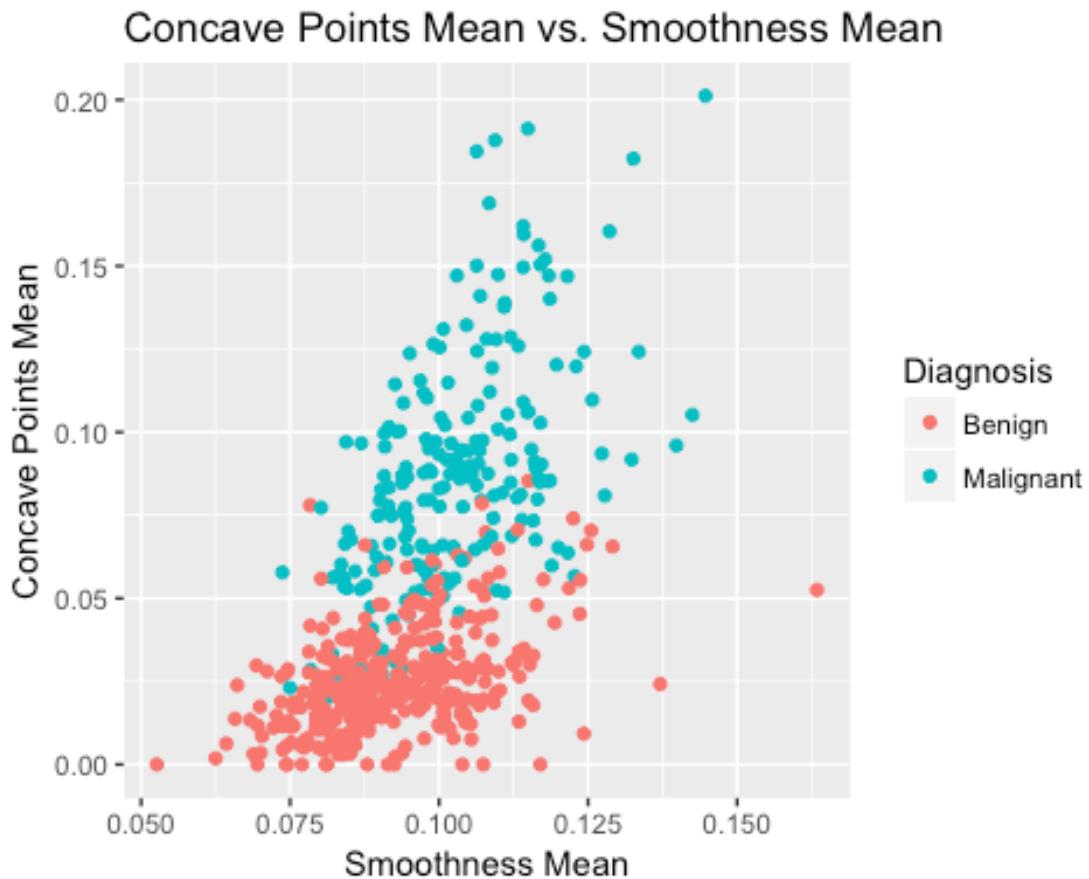


```
#concavity_mean vs. smoothness_mean
smoothness_concavity_mean_scatter <- ggplot(data = cancer, aes(x = smoothness_mean, y = concavity_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Mean", y = "Concavity Mean", title = "Concavity Mean vs. Smoothness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_concavity_mean_scatter
```

Concavity Mean vs. Smoothness Mean

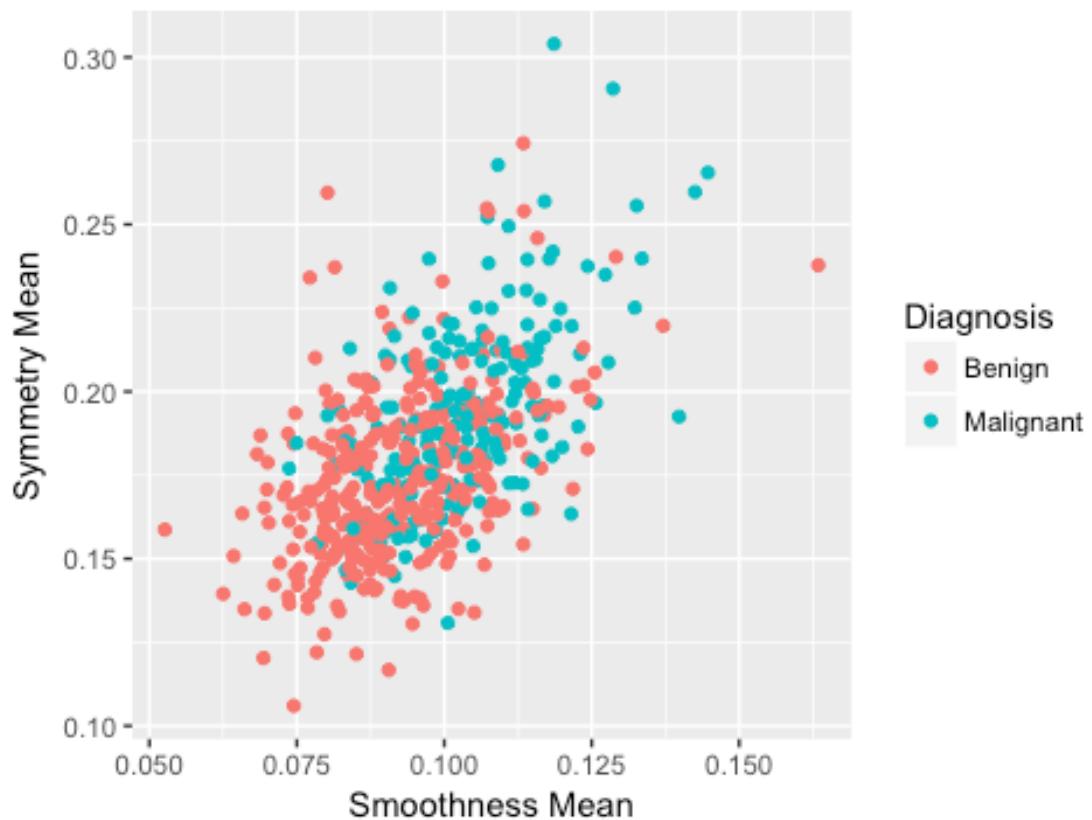


```
#concave.points_mean vs. smoothness_mean
smoothness_concave.points_mean_scatter <- ggplot(data = cancer, aes(x = smoothness_mean, y = concave.points_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Mean", y = "Concave Points Mean", title =
  "Concave Points Mean vs. Smoothness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_concave.points_mean_scatter
```



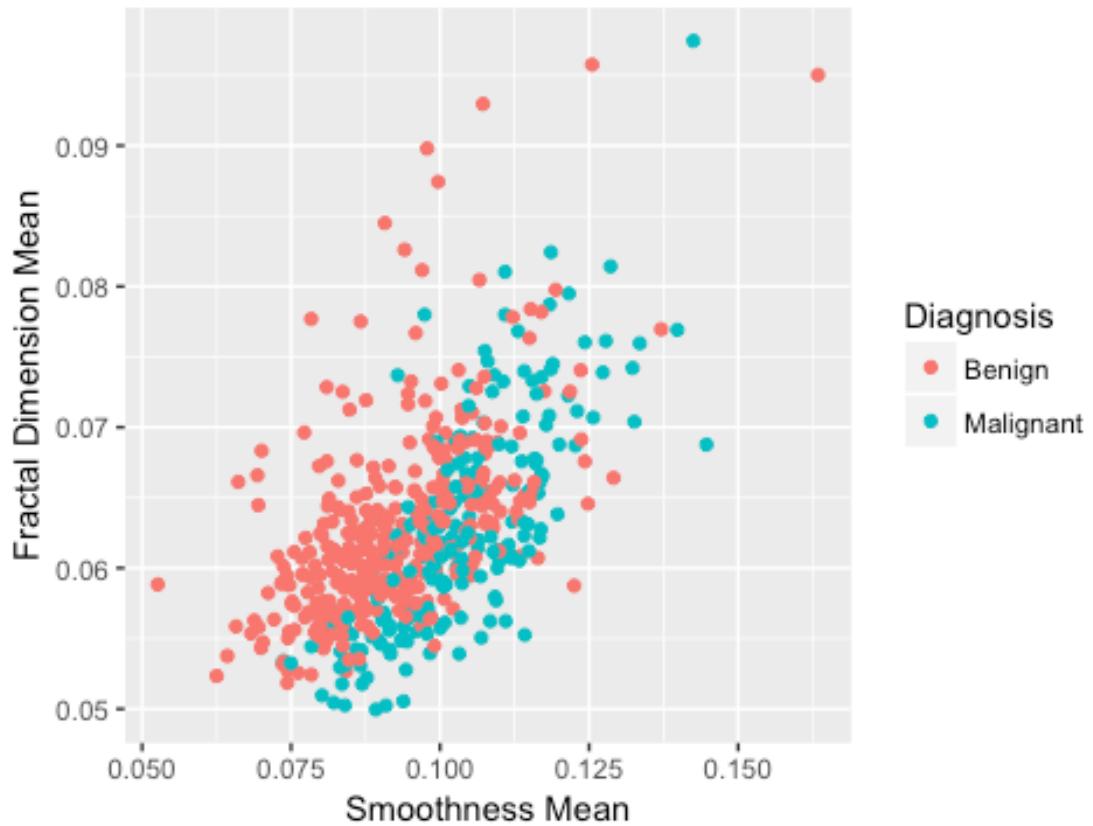
```
#symmetry_mean vs. smoothness_mean
smoothness_symmetry_mean_scatter <- ggplot(data = cancer, aes(x = smoothness_mean, y = symmetry_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Mean", y = "Symmetry Mean", title = "Symmetry Mean vs. Smoothness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_symmetry_mean_scatter
```

Symmetry Mean vs. Smoothness Mean



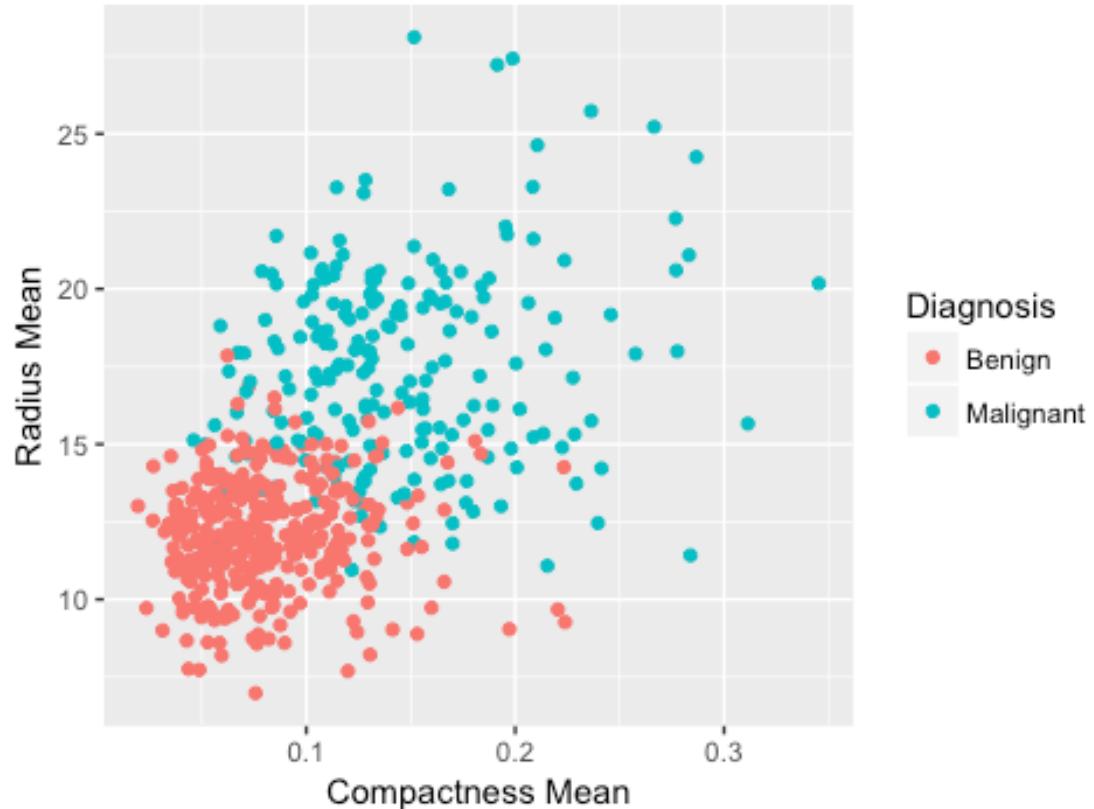
```
#fractal_dimension_mean vs. smoothness_mean
smoothness_fractal_dimension_mean_scatter <- ggplot(data = cancer, aes(x = smoothness_mean, y = fractal_dimension_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Mean", y = "Fractal Dimension Mean", title = "Fractal Dimension Mean vs. Smoothness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_fractal_dimension_mean_scatter
```

Fractal Dimension Mean vs. Smoothness Mean



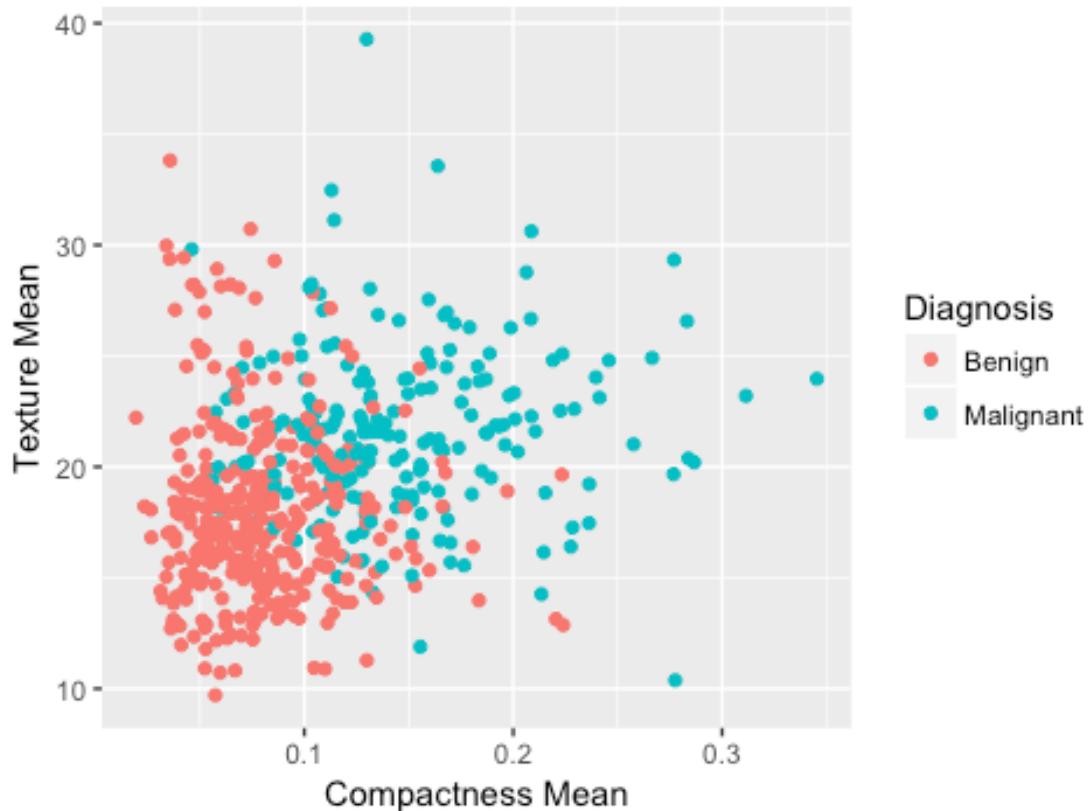
```
#radius_mean vs. compactness_mean
compactness_radius_mean_scatter <- ggplot(data = cancer, aes(x = compactness_mean, y = radius_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Mean", y = "Radius Mean", title = "Radius Mean vs. Compactness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_radius_mean_scatter
```

Radius Mean vs. Compactness Mean



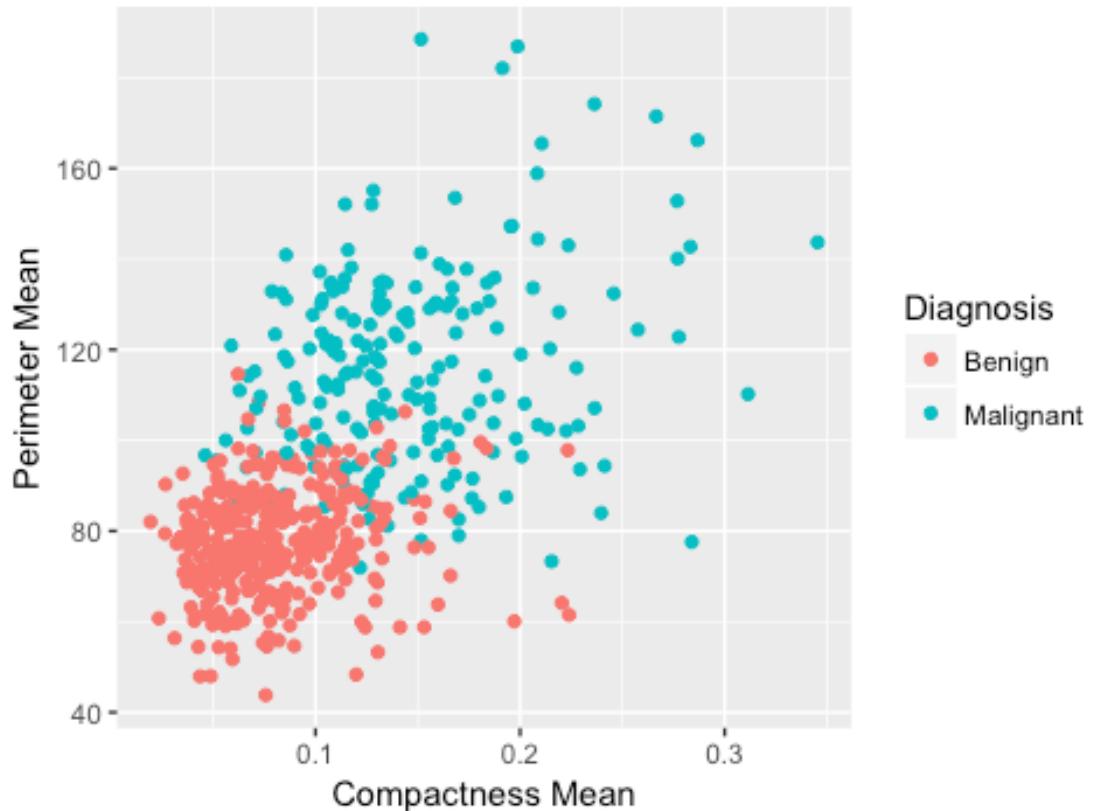
```
#texture_mean vs. compactness_mean
compactness_texture_mean_scatter <- ggplot(data = cancer, aes(x = compactness_mean, y = texture_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Mean", y = "Texture Mean", title = "Texture Mean vs. Compactness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_texture_mean_scatter
```

Texture Mean vs. Compactness Mean



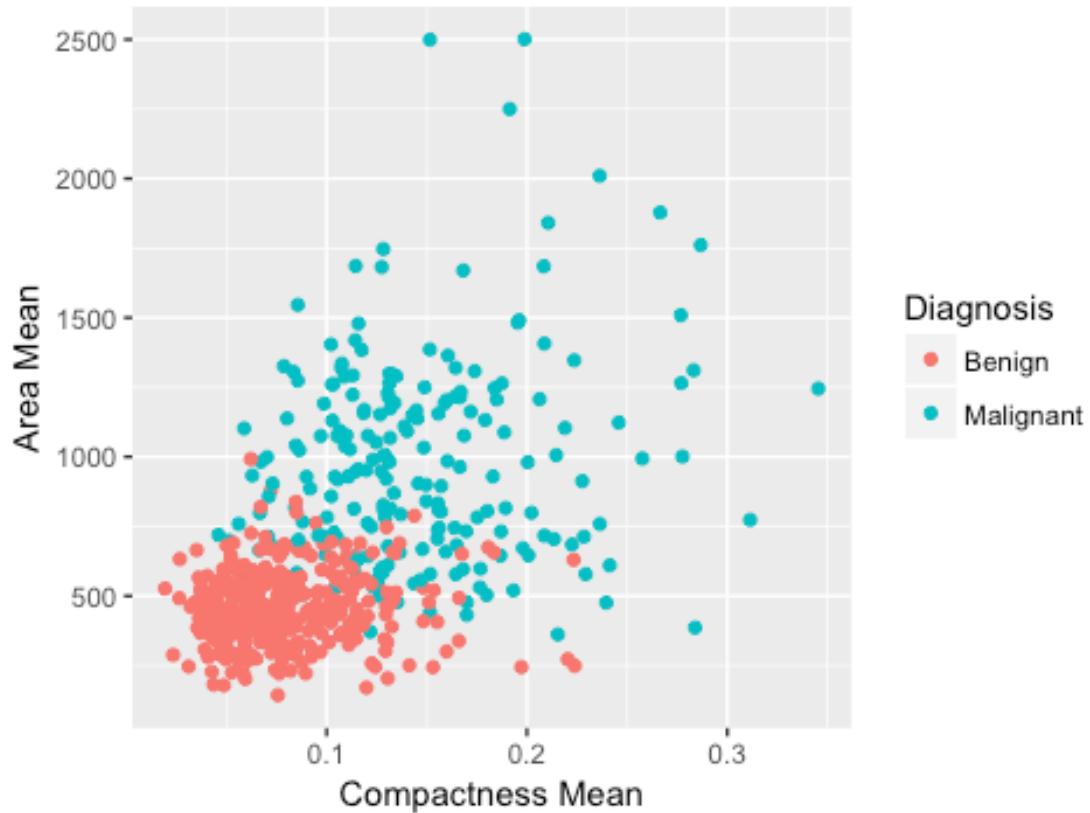
```
#perimeter_mean vs. compactness_mean
compactness_perimeter_mean_scatter <- ggplot(data = cancer, aes(x = compactness_mean, y = perimeter_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Mean", y = "Perimeter Mean", title = "Perimeter Mean vs. Compactness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_perimeter_mean_scatter
```

Perimeter Mean vs. Compactness Mean



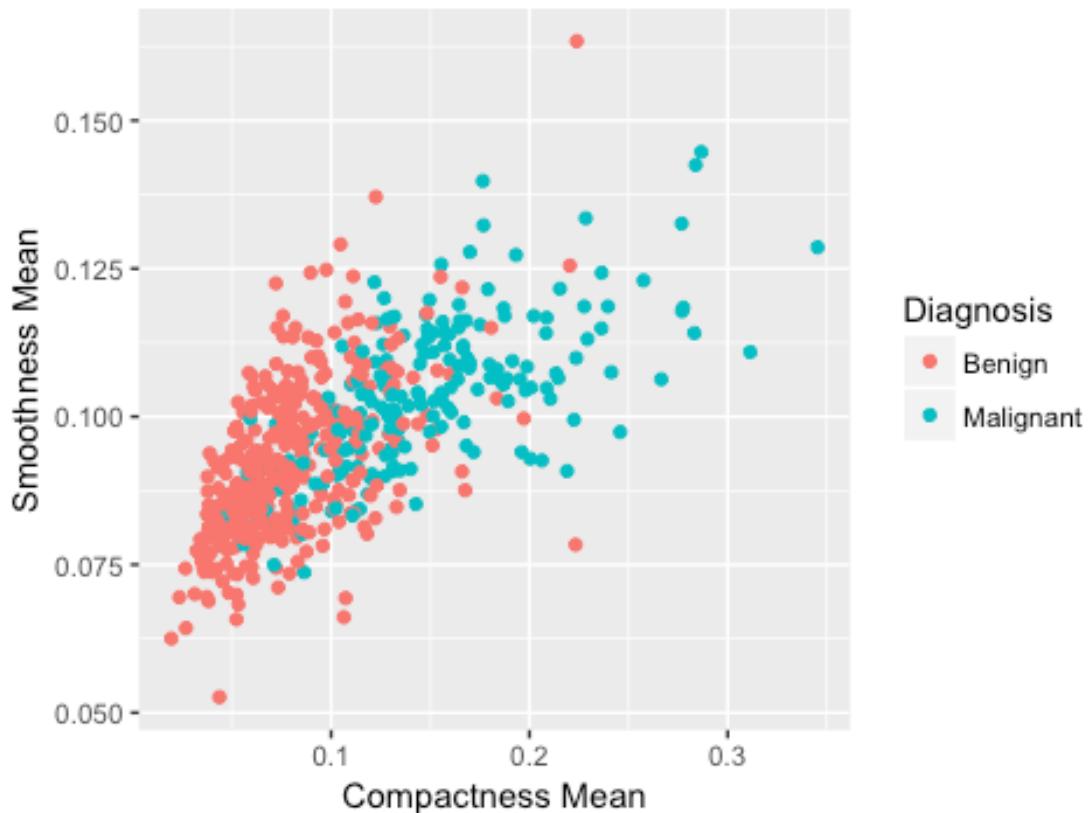
```
#area_mean vs. compactness_mean
compactness_area_mean_scatter <- ggplot(data = cancer, aes(x = compactness_mean, y = area_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Mean", y = "Area Mean", title = "Area Mean vs. Compactness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_area_mean_scatter
```

Area Mean vs. Compactness Mean



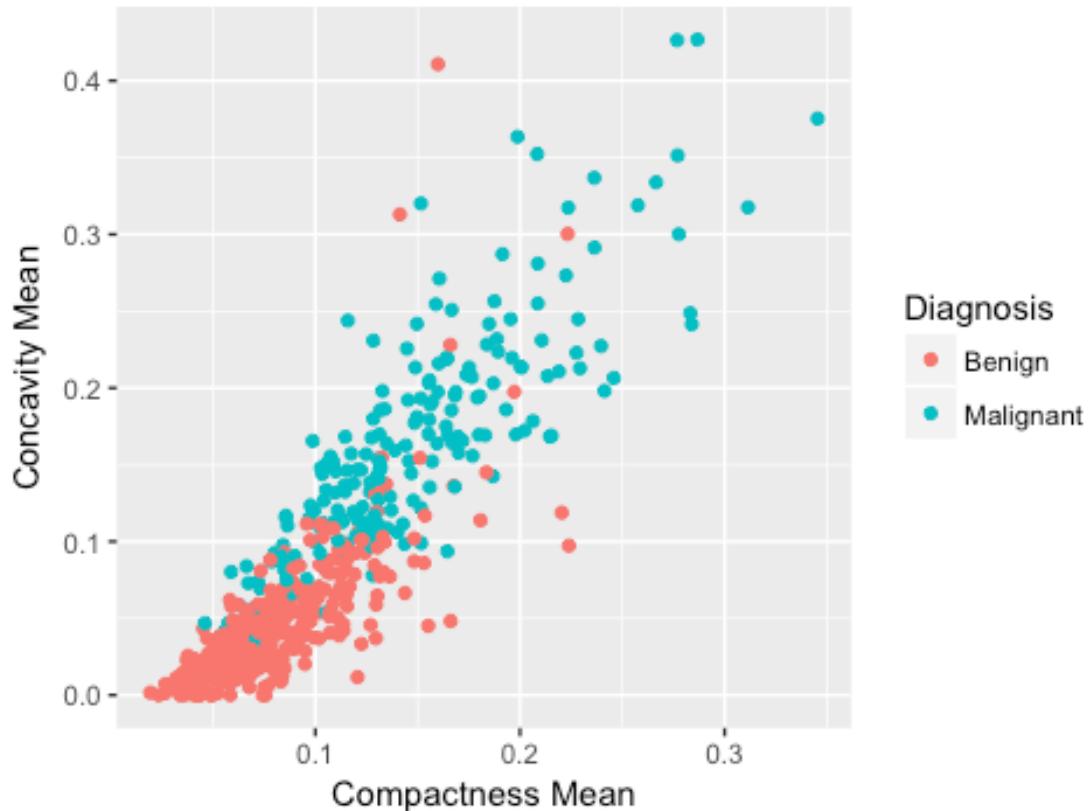
```
#smoothness_mean vs. compactness_mean
compactness_smoothness_mean_scatter <- ggplot(data = cancer, aes(x = compactness_mean, y = smoothness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Mean", y = "Smoothness Mean", title =
  "Smoothness Mean vs. Compactness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_smoothness_mean_scatter
```

Smoothness Mean vs. Compactness Mean



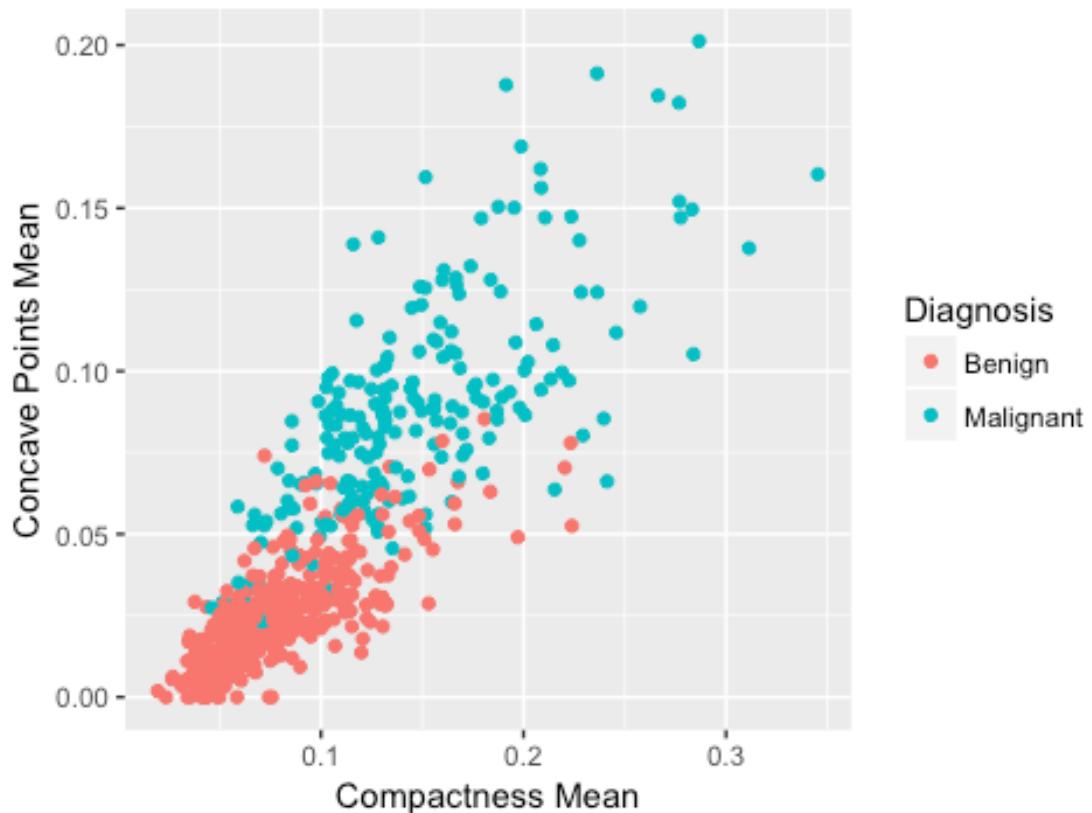
```
#concavity_mean vs. compactness_mean
compactness_concavity_mean_scatter <- ggplot(data = cancer, aes(x = compactness_mean, y = concavity_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Mean", y = "Concavity Mean", title = "Concavity Mean vs. Compactness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_concavity_mean_scatter
```

Concavity Mean vs. Compactness Mean



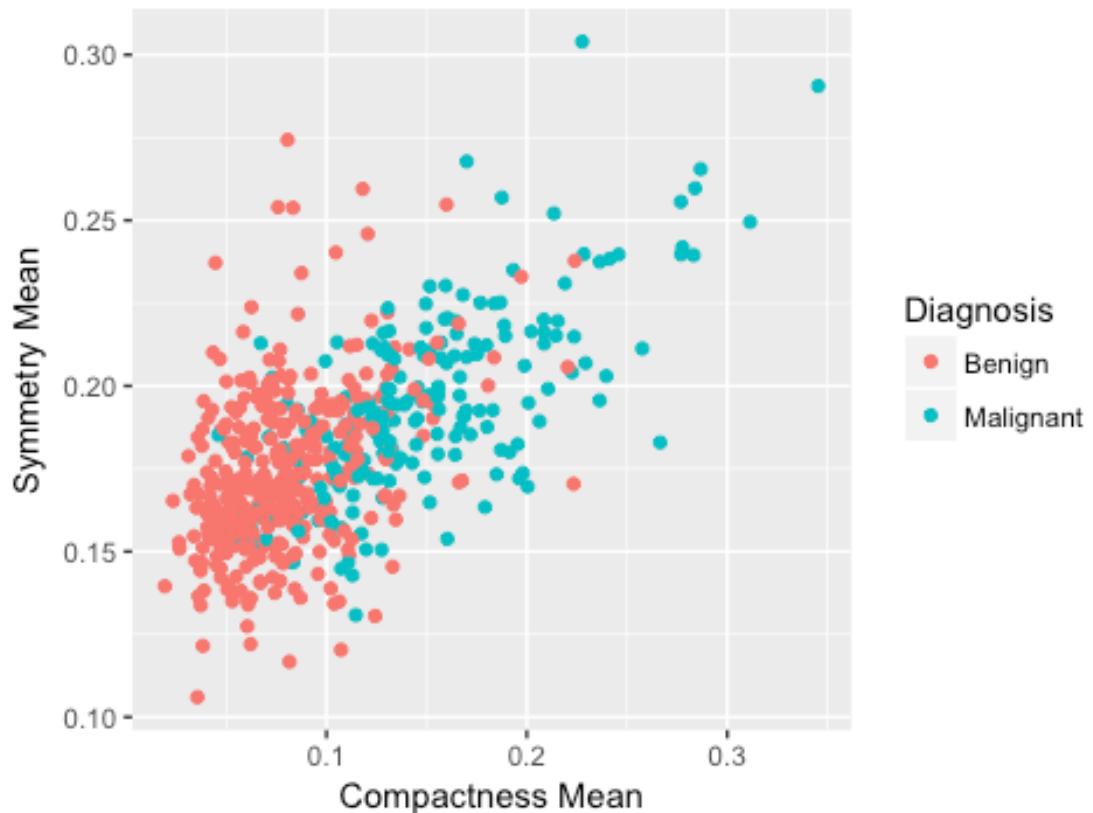
```
#concave.points_mean vs. compactness_mean
compactness_concave.points_mean_scatter <- ggplot(data = cancer, aes(x = compactness_mean, y = concave.points_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Mean", y = "Concave Points Mean", title = "Concave Points Mean vs. Compactness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_concave.points_mean_scatter
```

Concave Points Mean vs. Compactness Mean



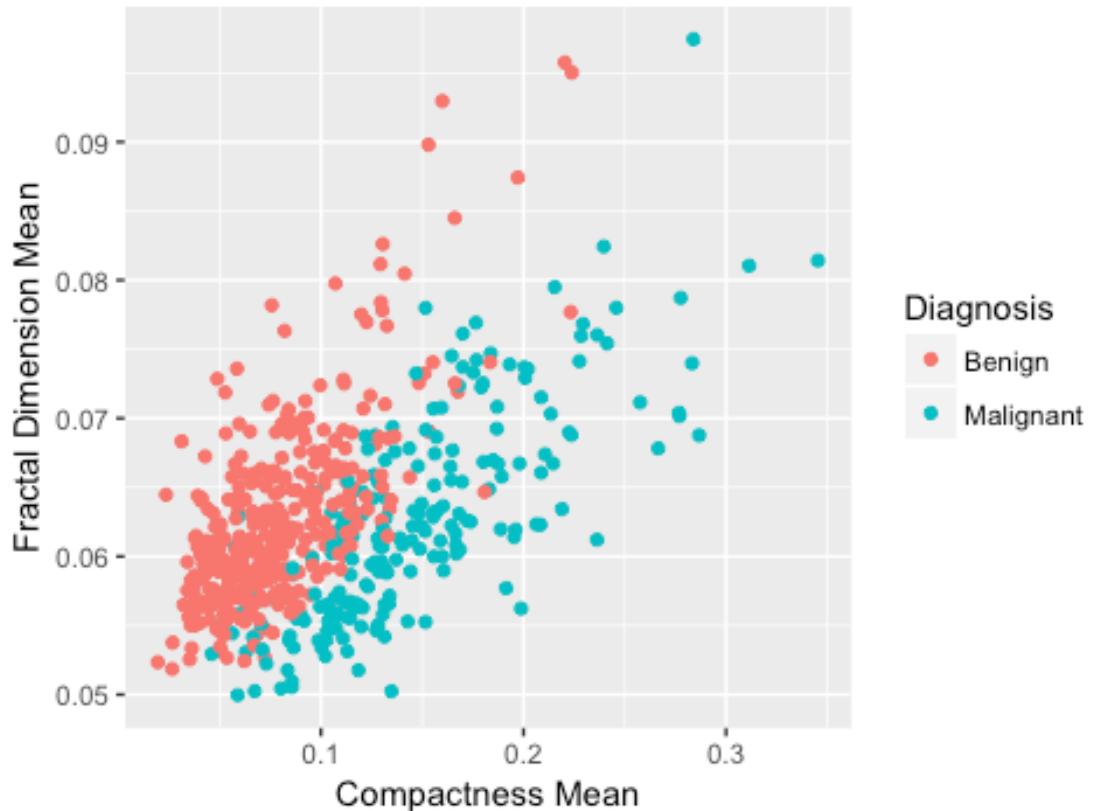
```
#symmetry_mean vs. compactness_mean
compactness_symmetry_mean_scatter <- ggplot(data = cancer, aes(x = compactness_mean, y = symmetry_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Mean", y = "Symmetry Mean", title = "Symmetry Mean vs. Compactness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_symmetry_mean_scatter
```

Symmetry Mean vs. Compactness Mean



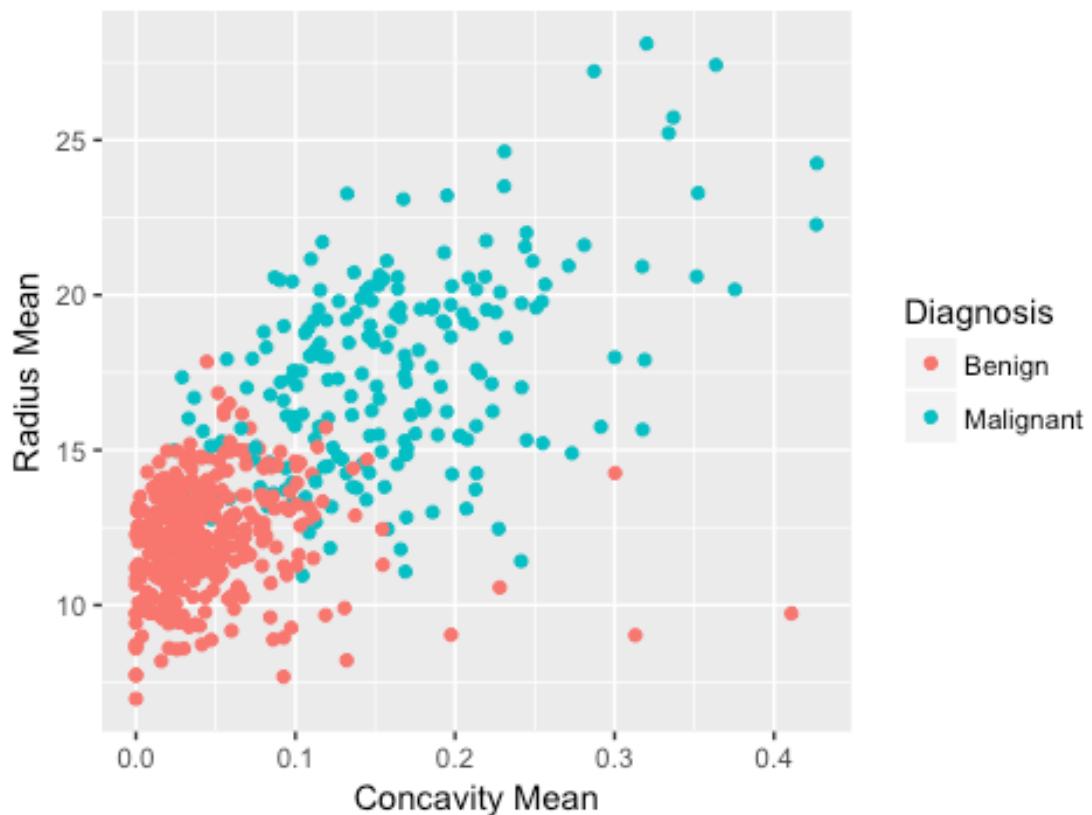
```
#fractal_dimension_mean vs. compactness_mean
compactness_fractal_dimension_mean_scatter <- ggplot(data = cancer, aes(x = compactness_mean, y = fractal_dimension_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Mean", y = "Fractal Dimension Mean", title = "Fractal Dimension Mean vs. Compactness Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_fractal_dimension_mean_scatter
```

Fractal Dimension Mean vs. Compactness Mean



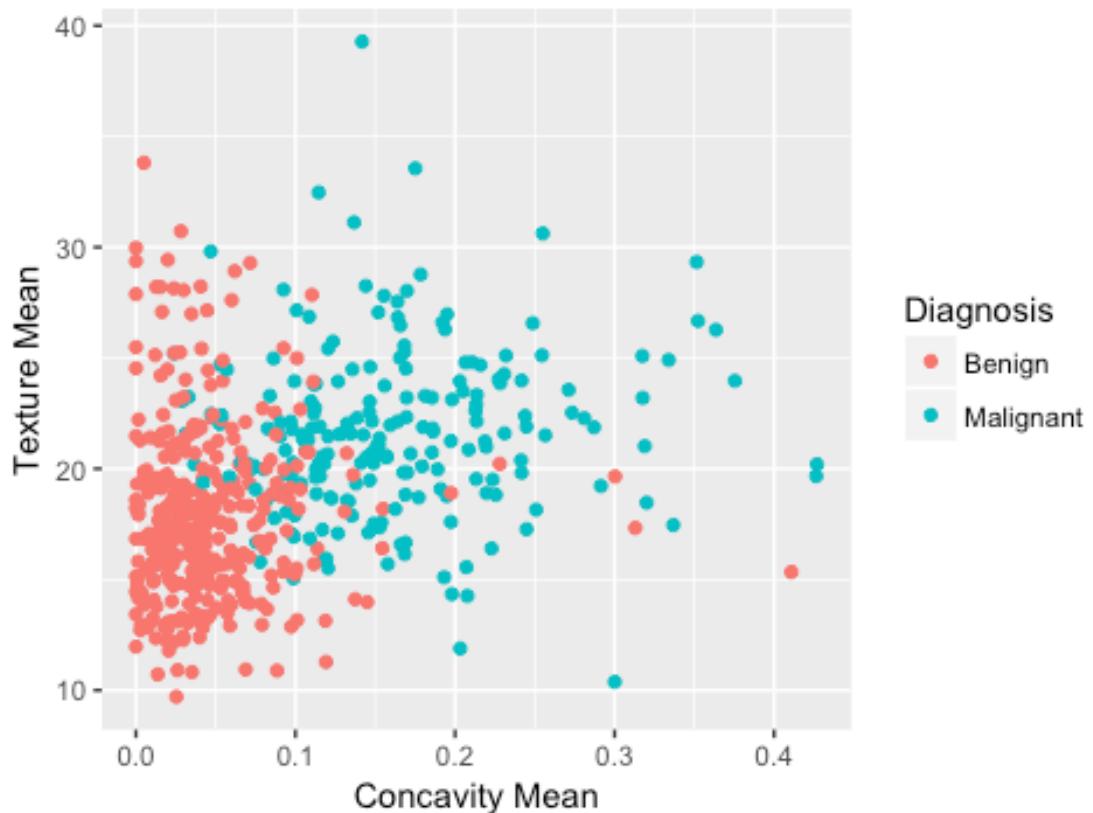
```
#radius_mean vs. concavity_mean
concavity_radius_mean_scatter <- ggplot(data = cancer, aes(x = concavity_mean,
, y = radius_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Mean", y = "Radius Mean", title = "Radius Mean vs. Concavity Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_radius_mean_scatter
```

Radius Mean vs. Concavity Mean



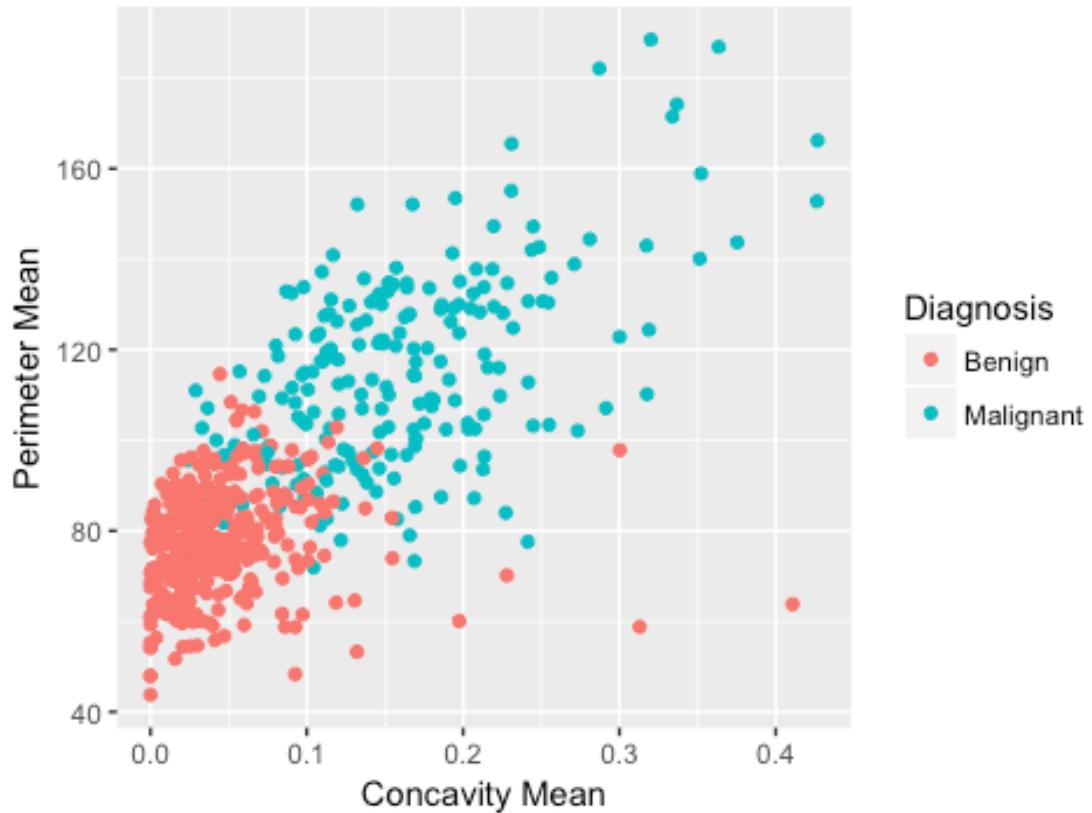
```
#texture_mean vs. concavity_mean
concavity_texture_mean_scatter <- ggplot(data = cancer, aes(x = concavity_mean, y = texture_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Mean", y = "Texture Mean", title = "Texture Mean vs. Concavity Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_texture_mean_scatter
```

Texture Mean vs. Concavity Mean



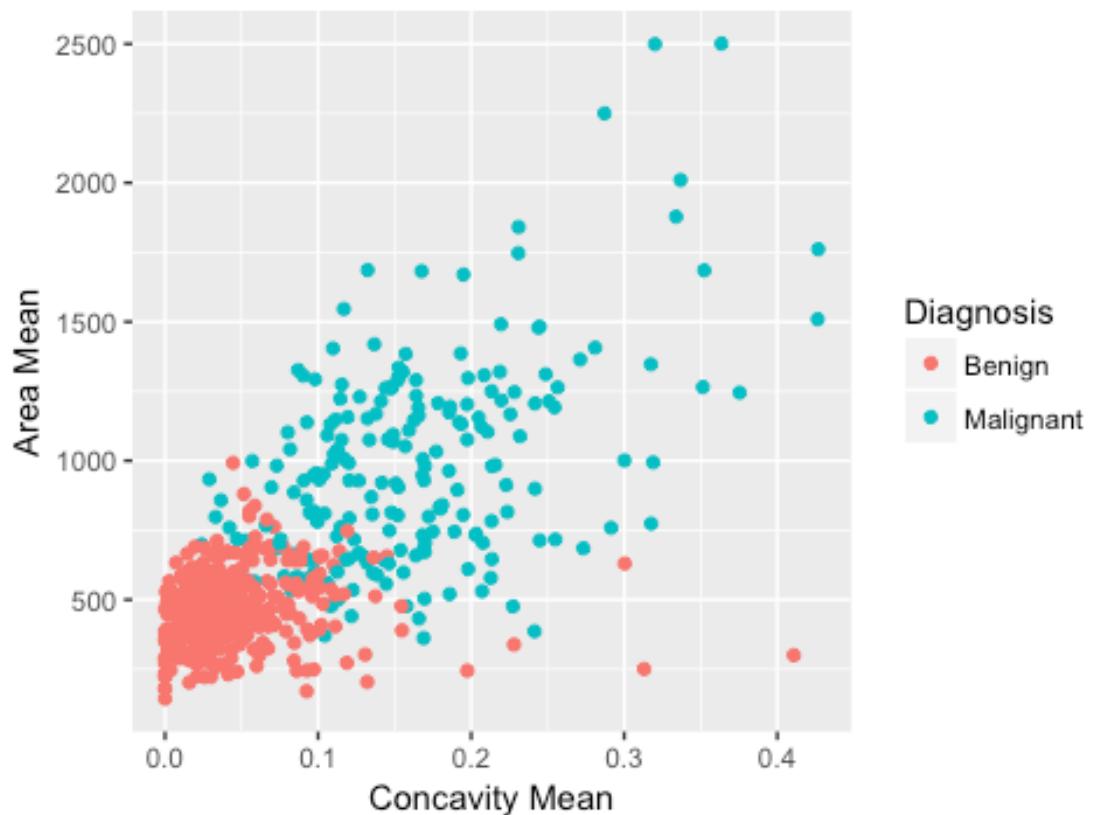
```
#perimeter_mean vs. concavity_mean
concavity_perimeter_mean_scatter <- ggplot(data = cancer, aes(x = concavity_mean, y = perimeter_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Mean", y = "Perimeter Mean", title = "Perimeter Mean vs. Concavity Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_perimeter_mean_scatter
```

Perimeter Mean vs. Concavity Mean



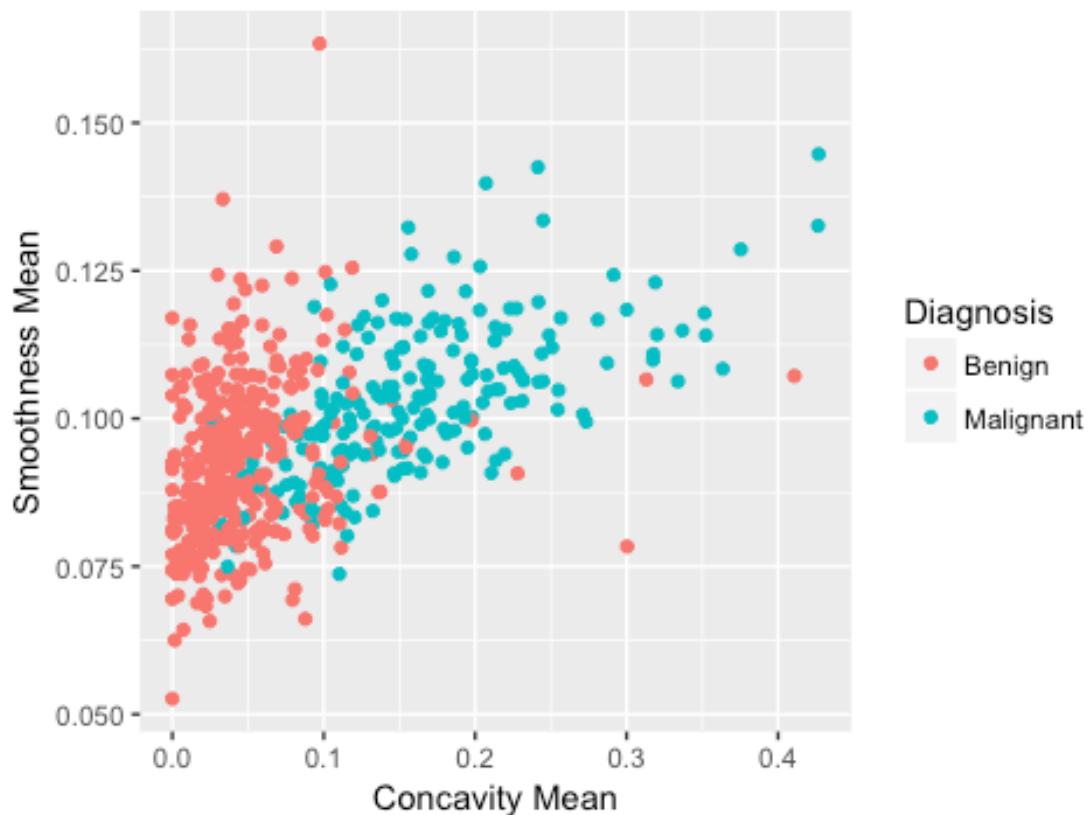
```
#area_mean vs. concavity_mean
concavity_area_mean_scatter <- ggplot(data = cancer, aes(x = concavity_mean,
y = area_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Mean", y = "Area Mean", title = "Area Mean vs. Concavity Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_area_mean_scatter
```

Area Mean vs. Concavity Mean



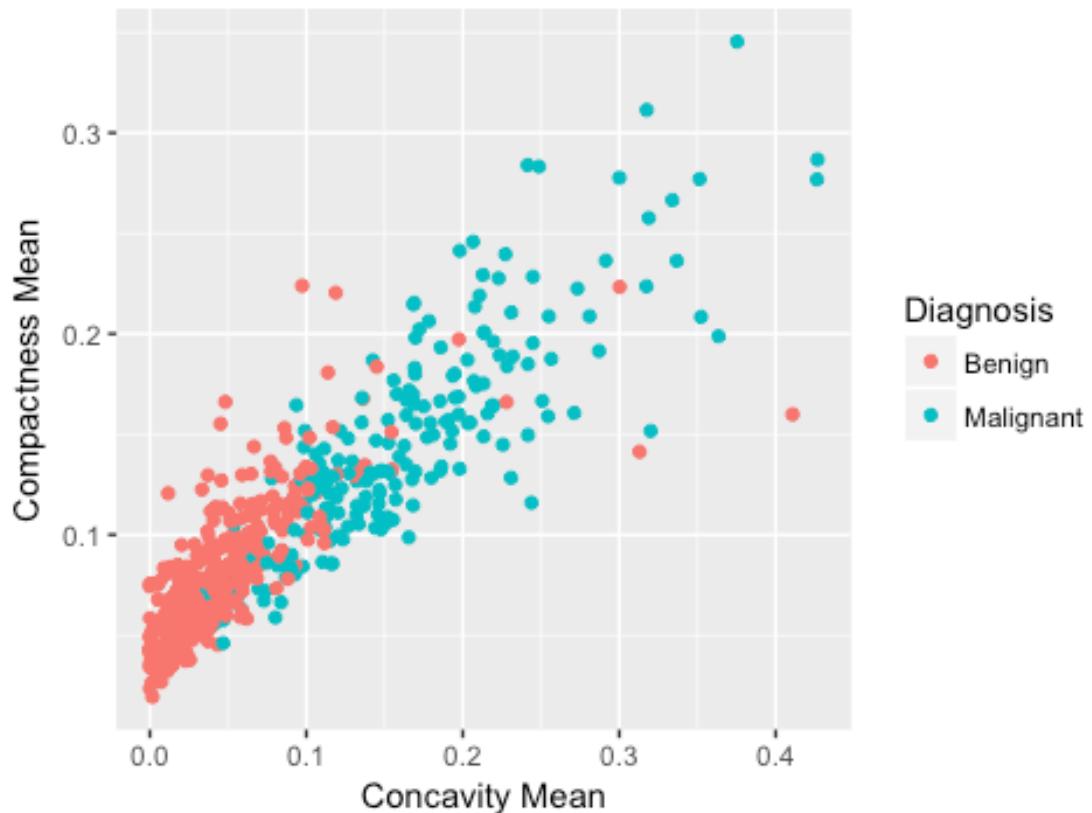
```
#smoothness_mean vs. concavity_mean
concavity_smoothness_mean_scatter <- ggplot(data = cancer, aes(x = concavity_mean, y = smoothness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Mean", y = "Smoothness Mean", title = "Smoothness Mean vs. Concavity Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_smoothness_mean_scatter
```

Smoothness Mean vs. Concavity Mean



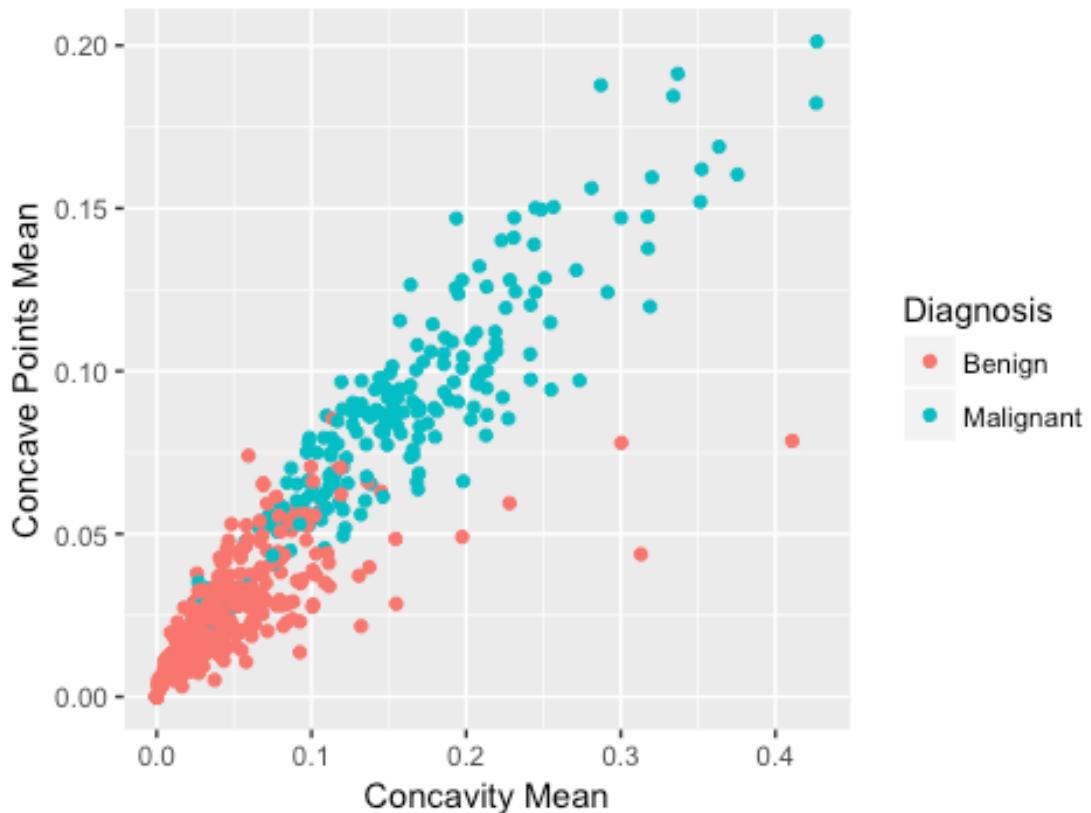
```
#compactness_mean vs. concavity_mean
concavity_compactness_mean_scatter <- ggplot(data = cancer, aes(x = concavity_mean, y = compactness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Mean", y = "Compactness Mean", title = "Compactness Mean vs. Concavity Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_compactness_mean_scatter
```

Compactness Mean vs. Concavity Mean



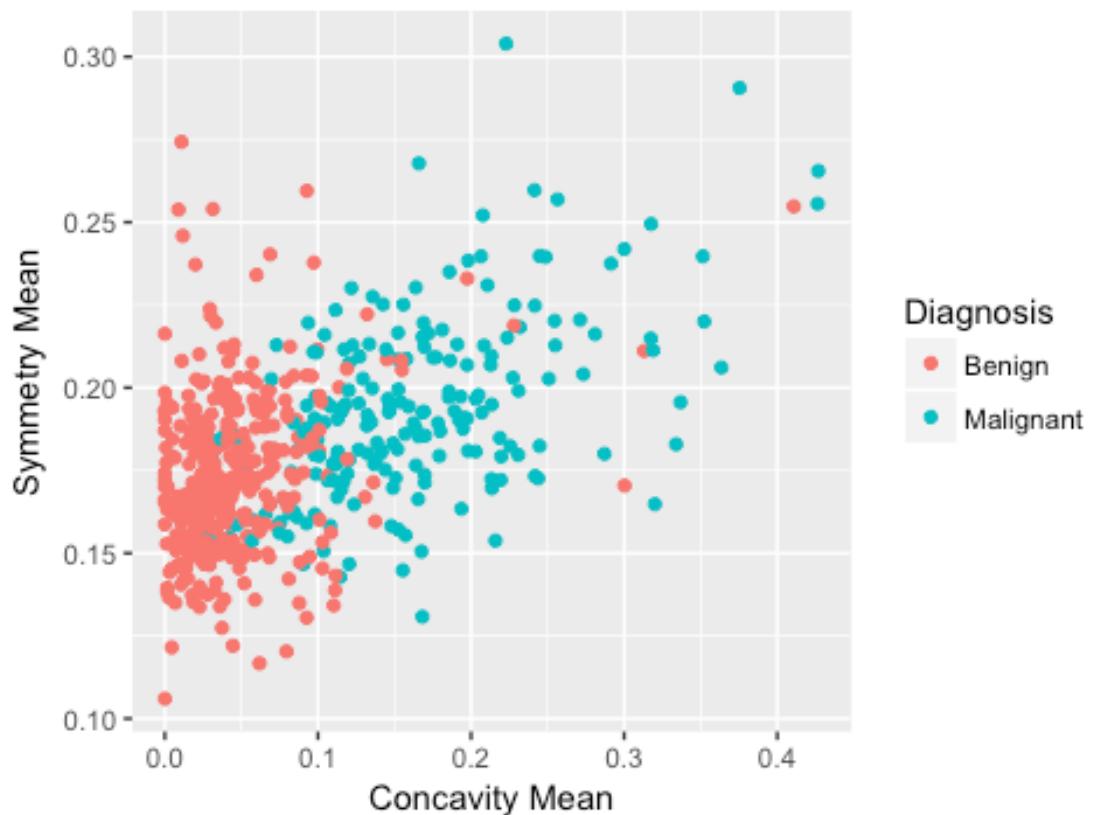
```
#concave.points_mean vs. concavity_mean
concavity_concave.points_mean_scatter <- ggplot(data = cancer, aes(x = concavity_mean, y = concave.points_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Mean", y = "Concave Points Mean", title =
  "Concave Points Mean vs. Concavity Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_concave.points_mean_scatter
```

Concave Points Mean vs. Concavity Mean



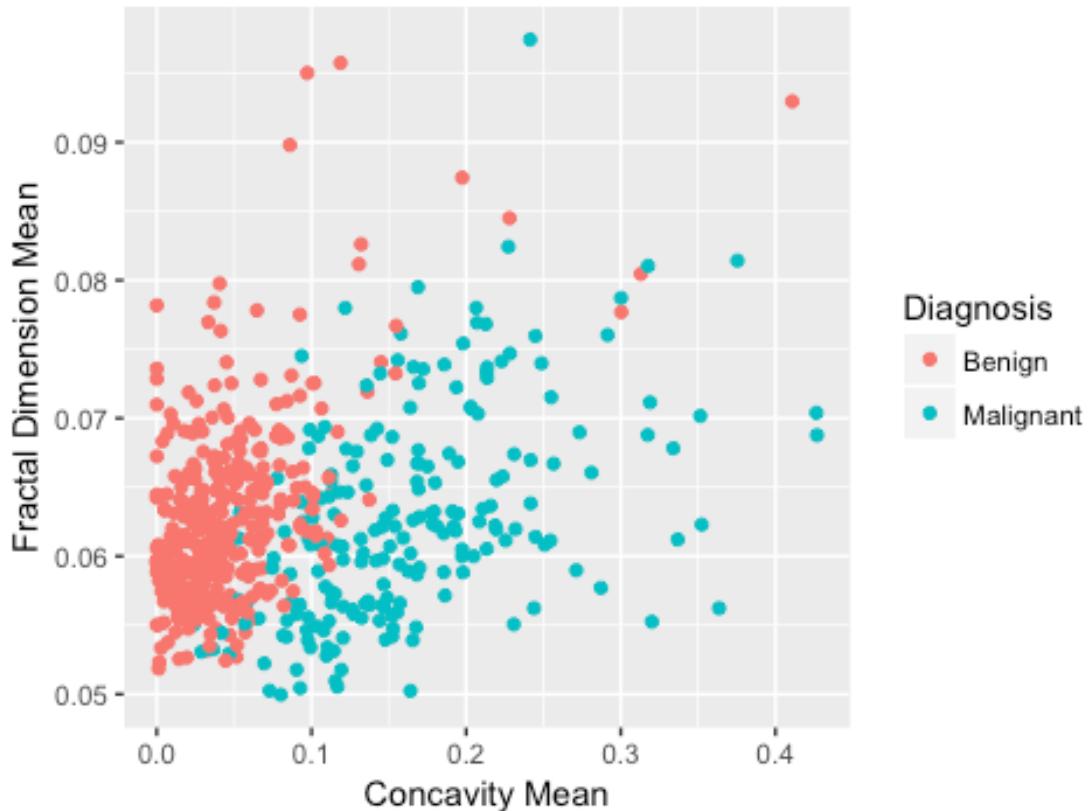
```
#symmetry_mean vs. concavity_mean
concavity_symmetry_mean_scatter <- ggplot(data = cancer, aes(x = concavity_mean, y = symmetry_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Mean", y = "Symmetry Mean", title = "Symmetry Mean vs. Concavity Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_symmetry_mean_scatter
```

Symmetry Mean vs. Concavity Mean



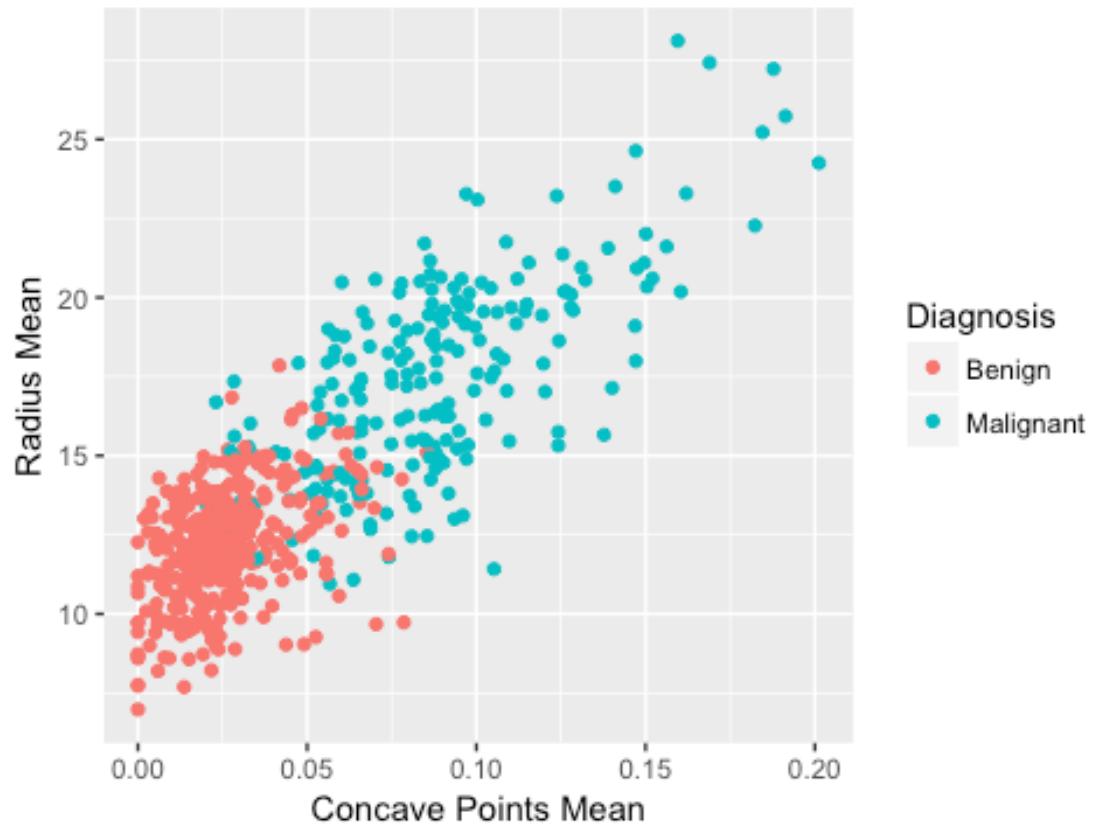
```
#fractal_dimension_mean vs. concavity_mean
concavity_fractal_dimension_mean_scatter <- ggplot(data = cancer, aes(x = concavity_mean, y = fractal_dimension_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Mean", y = "Fractal Dimension Mean", title = "Fractal Dimension Mean vs. Concavity Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_fractal_dimension_mean_scatter
```

Fractal Dimension Mean vs. Concavity Mean



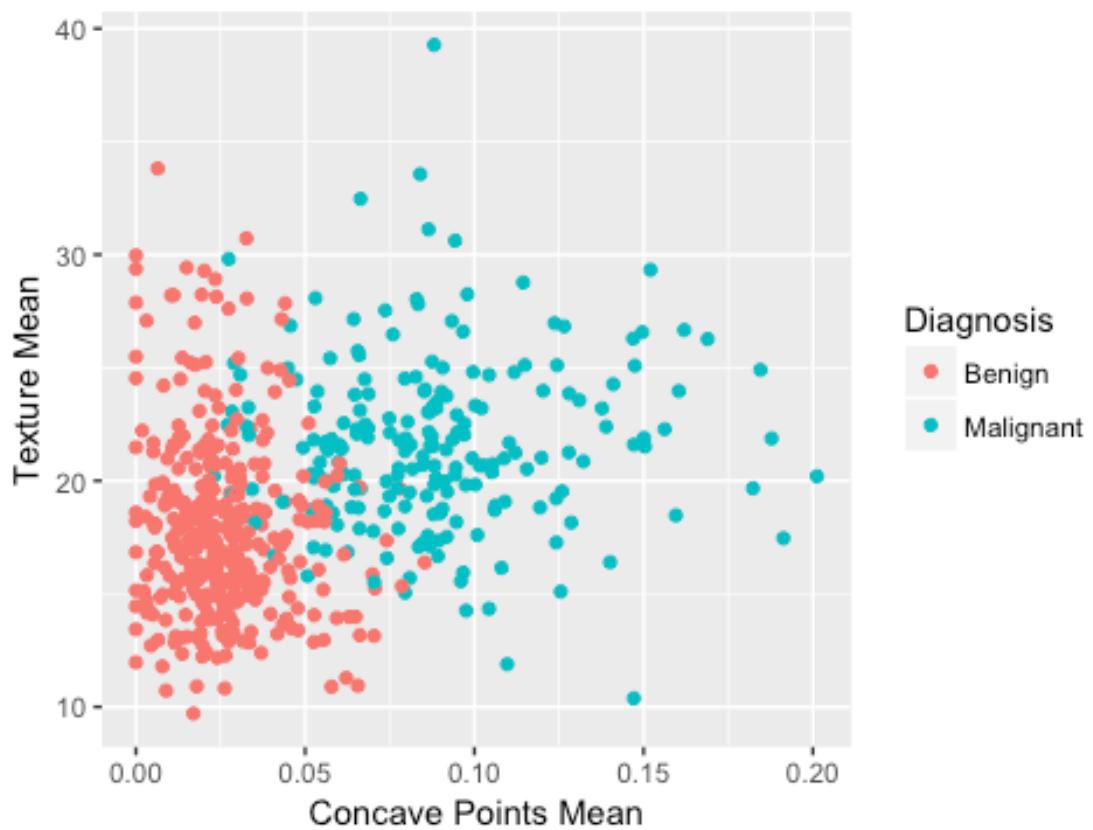
```
#radius_mean vs. concave.points_mean
concave.points_radius_mean_scatter <- ggplot(data = cancer, aes(x = concave.points_mean, y = radius_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Mean", y = "Radius Mean", title = "Radius Mean vs. Concave Points Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_radius_mean_scatter
```

Radius Mean vs. Concave Points Mean



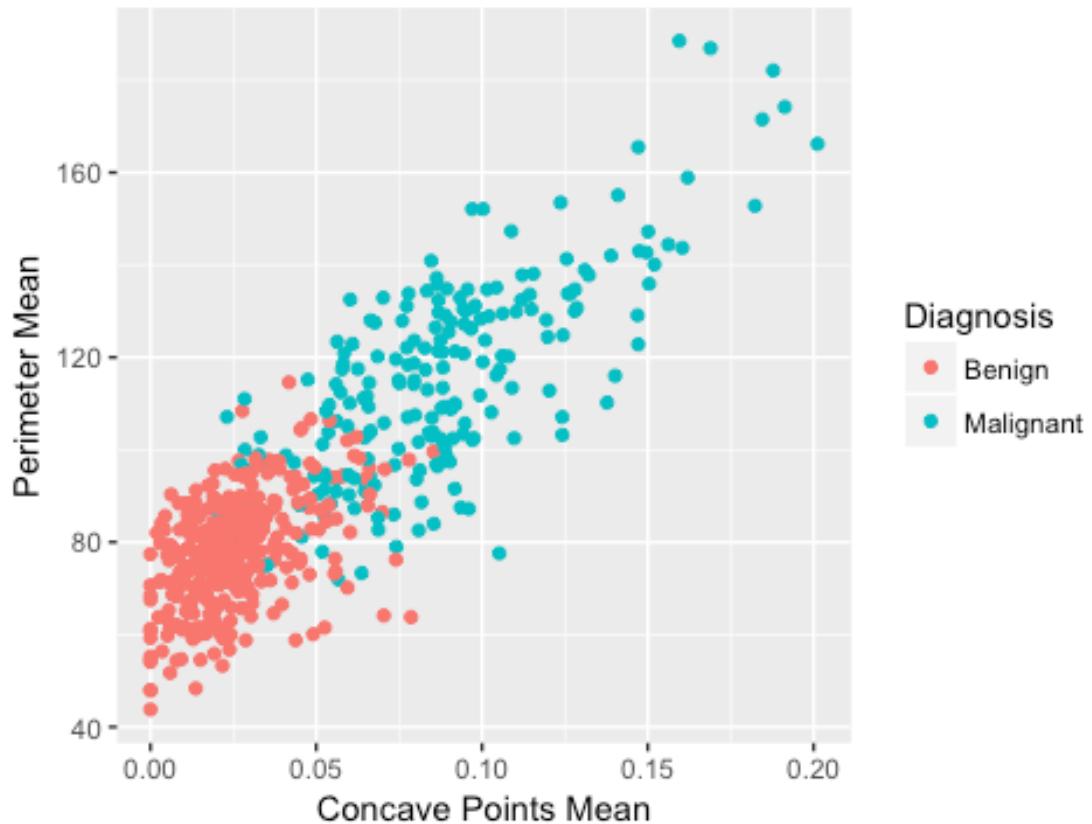
```
#texture_mean vs. concave.points_mean
concave.points_texture_mean_scatter <- ggplot(data = cancer, aes(x = concave.
points_mean, y = texture_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Mean", y = "Texture Mean", title =
"Texture Mean vs. Concave Points Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_texture_mean_scatter
```

Texture Mean vs. Concave Points Mean



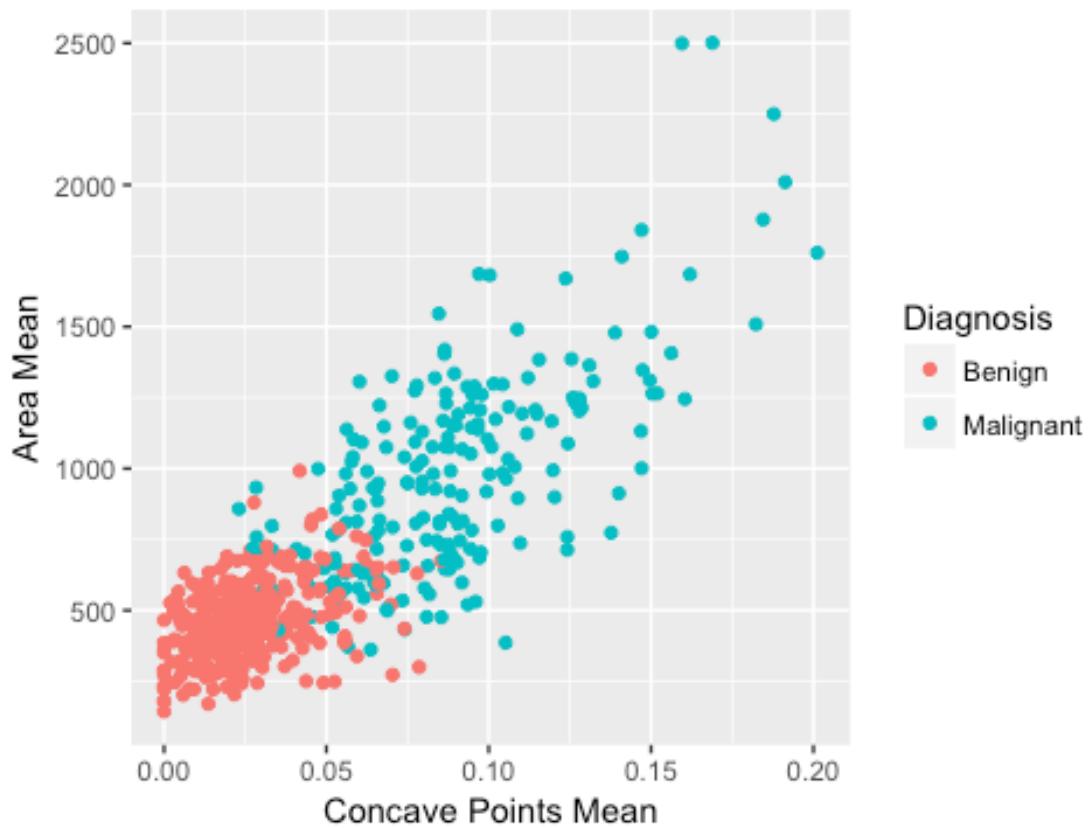
```
#perimeter_mean vs. concave.points_mean
concave.points_perimeter_mean_scatter <- ggplot(data = cancer, aes(x = concave.points_mean, y = perimeter_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Mean", y = "Perimeter Mean", title =
  "Perimeter Mean vs. Concave Points Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_perimeter_mean_scatter
```

Perimeter Mean vs. Concave Points Mean



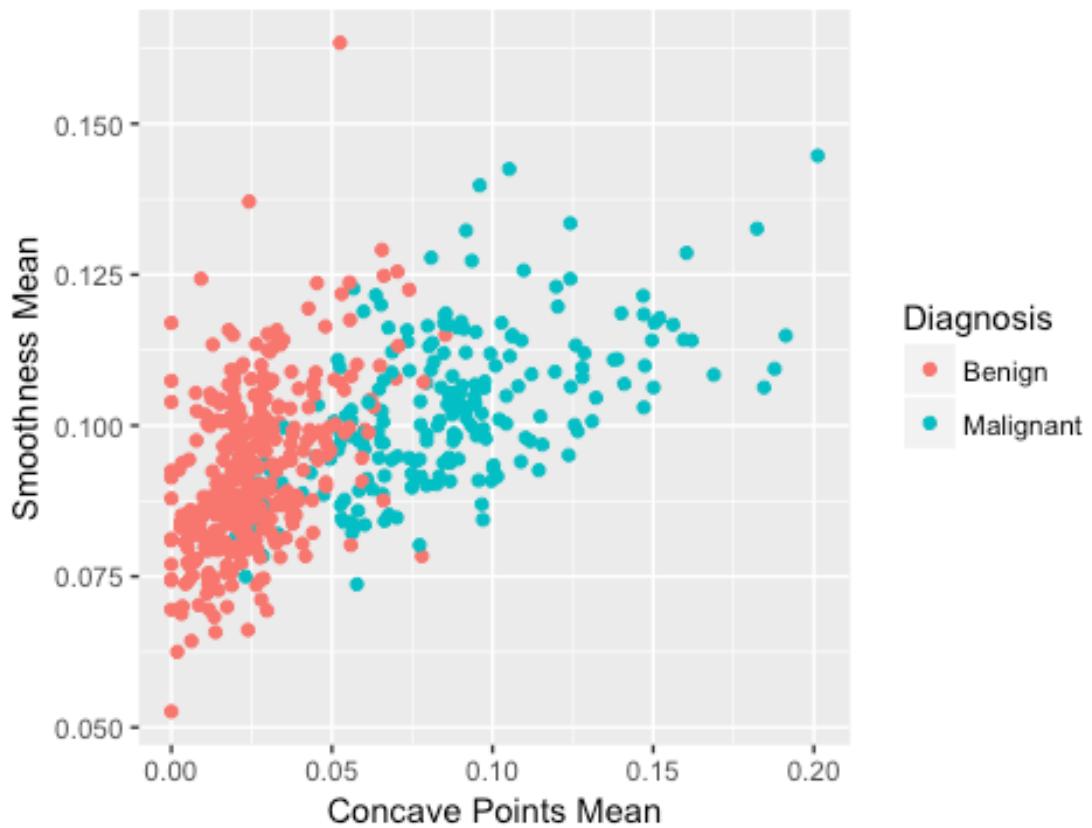
```
#area_mean vs. concave.points_mean
concave.points_area_mean_scatter <- ggplot(data = cancer, aes(x = concave.points_mean, y = area_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Mean", y = "Area Mean", title = "Area Mean vs. Concave Points Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_area_mean_scatter
```

Area Mean vs. Concave Points Mean



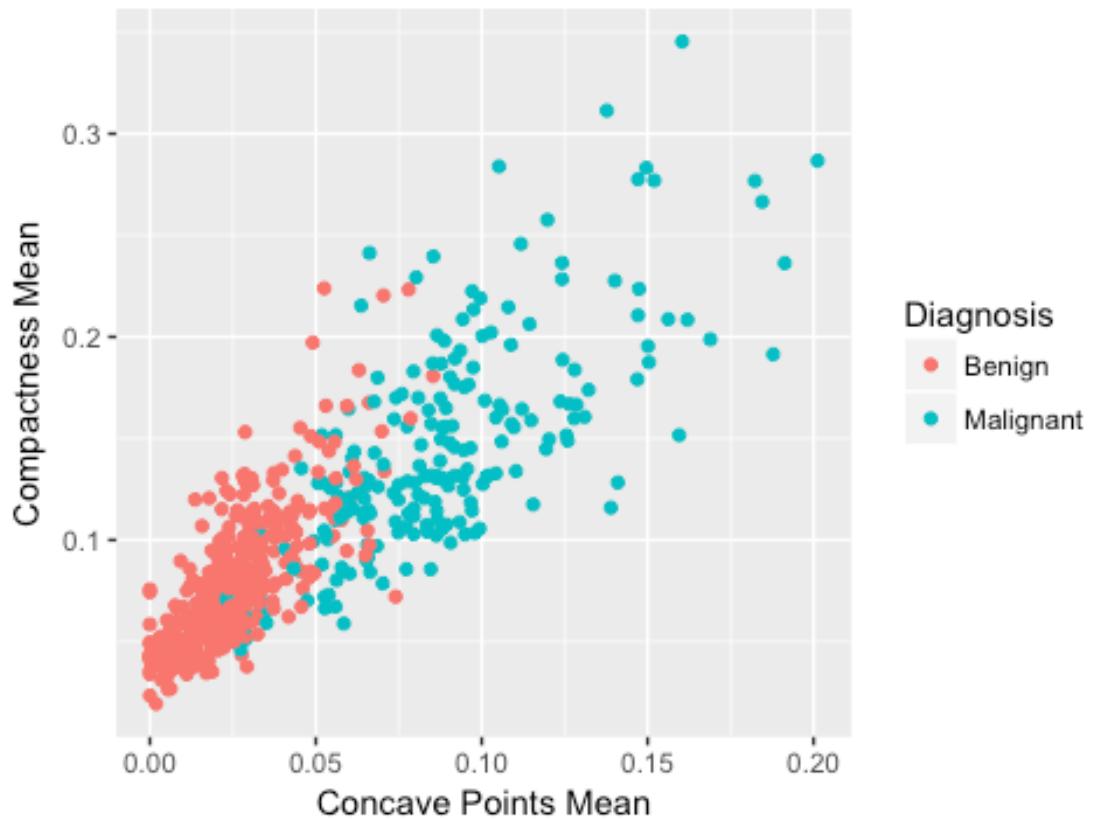
```
#smoothness_mean vs. concave.points_mean
concave.points_smoothness_mean_scatter <- ggplot(data = cancer, aes(x = concave.points_mean, y = smoothness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Mean", y = "Smoothness Mean", title =
  "Smoothness Mean vs. Concave Points Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_smoothness_mean_scatter
```

Smoothness Mean vs. Concave Points Mean



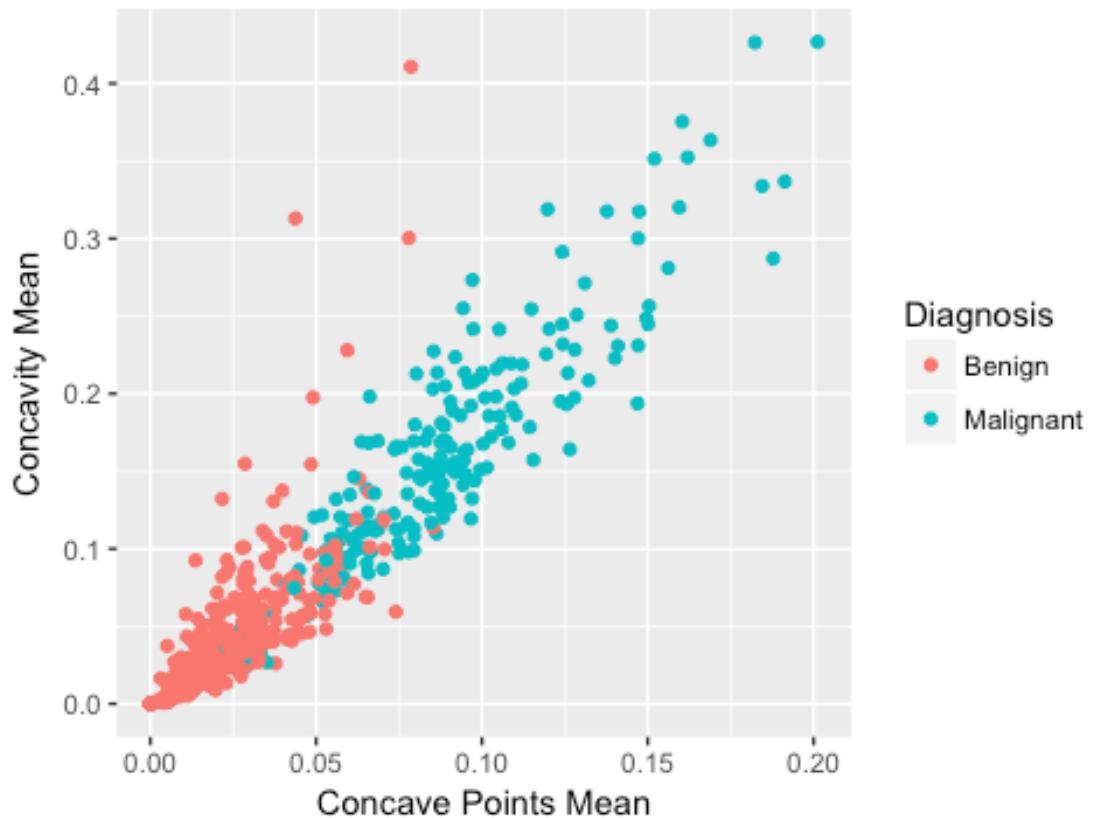
```
#compactness_mean vs. concave.points_mean
concave.points_compactness_mean_scatter <- ggplot(data = cancer, aes(x = concave.points_mean, y = compactness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Mean", y = "Compactness Mean", title = "Compactness Mean vs. Concave Points Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_compactness_mean_scatter
```

Compactness Mean vs. Concave Points Mean



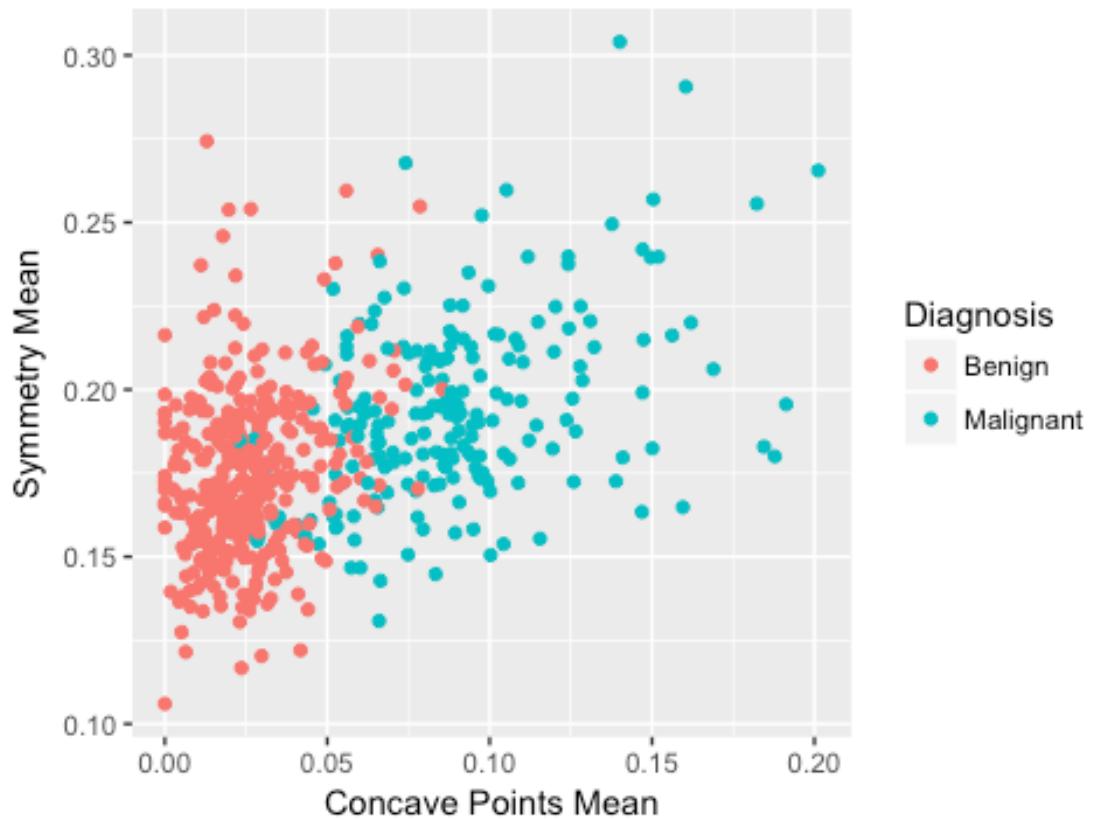
```
#concavity_mean vs. concave.points_mean
concave.points_concavity_mean_scatter <- ggplot(data = cancer, aes(x = concave.points_mean, y = concavity_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Mean", y = "Concavity Mean", title =
  "Concavity Mean vs. Concave Points Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_concavity_mean_scatter
```

Concavity Mean vs. Concave Points Mean



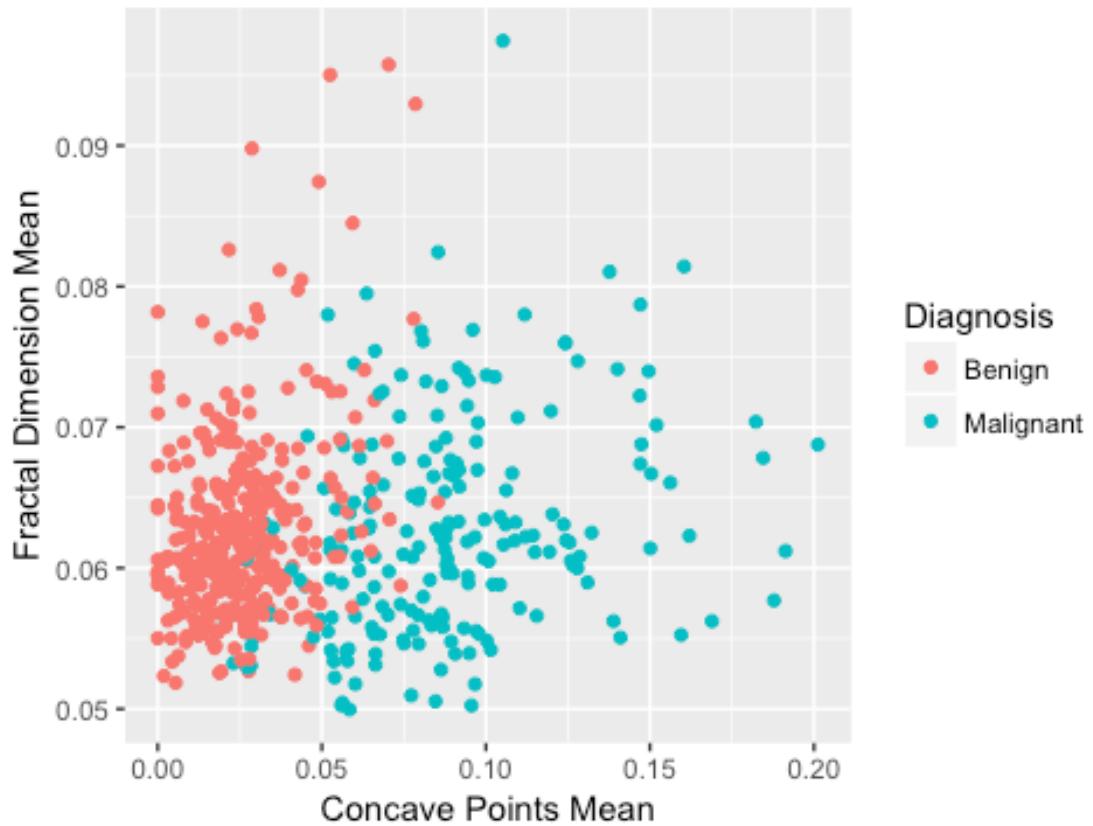
```
#symmetry_mean vs. concave.points_mean
concave.points_symmetry_mean_scatter <- ggplot(data = cancer, aes(x = concave
  .points_mean, y = symmetry_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Mean", y = "Symmetry Mean", title =
  "Symmetry Mean vs. Concave Points Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_symmetry_mean_scatter
```

Symmetry Mean vs. Concave Points Mean



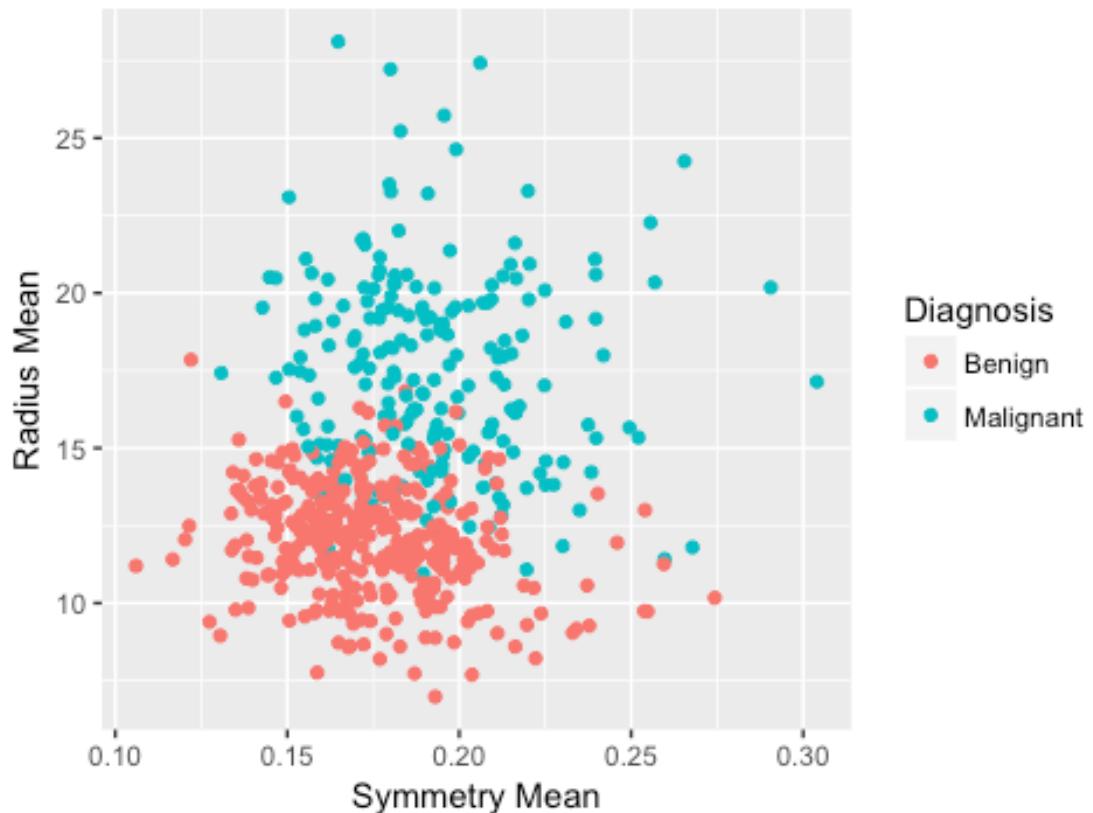
```
#fractal_dimension_mean vs. concave.points_mean
concave.points_fractal_dimension_mean_scatter <- ggplot(data = cancer, aes(x = concave.points_mean, y = fractal_dimension_mean, color = as.factor(diagnoses))) +
  geom_point() + labs(x = "Concave Points Mean", y = "Fractal Dimension Mean",
, title = "Fractal Dimension Mean vs. Concave Points Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_fractal_dimension_mean_scatter
```

Fractal Dimension Mean vs. Concave Points Mean



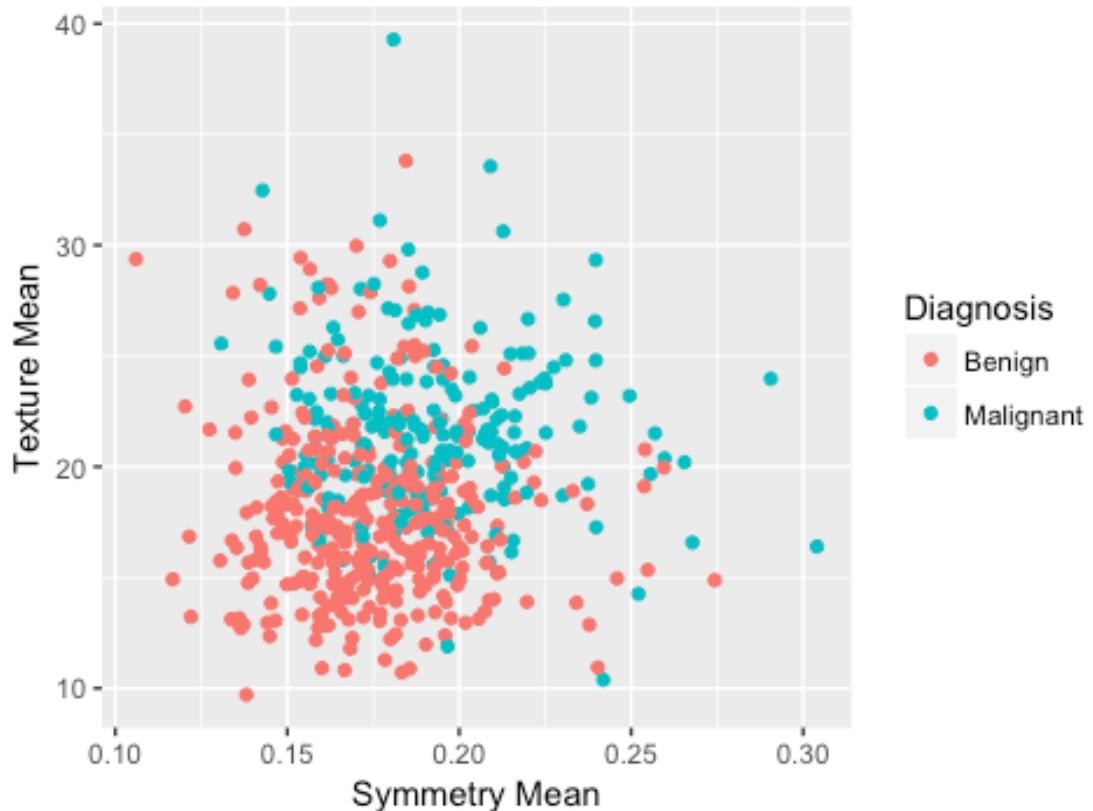
```
#radius_mean vs. symmetry_mean
symmetry_radius_mean_scatter <- ggplot(data = cancer, aes(x = symmetry_mean,
y = radius_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Mean", y = "Radius Mean", title = "Radius
Mean vs. Symmetry Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_radius_mean_scatter
```

Radius Mean vs. Symmetry Mean



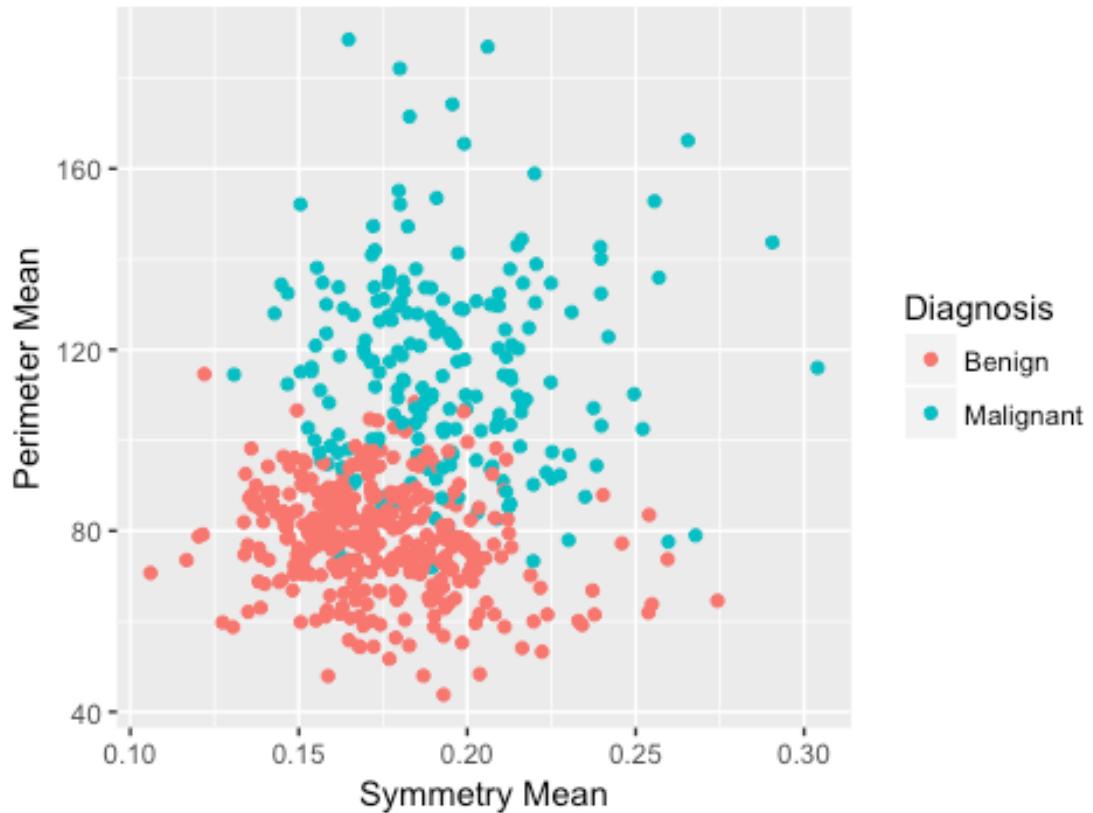
```
#texture_mean vs. symmetry_mean
symmetry_texture_mean_scatter <- ggplot(data = cancer, aes(x = symmetry_mean,
y = texture_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Mean", y = "Texture Mean", title = "Texture Mean vs. Symmetry Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_texture_mean_scatter
```

Texture Mean vs. Symmetry Mean



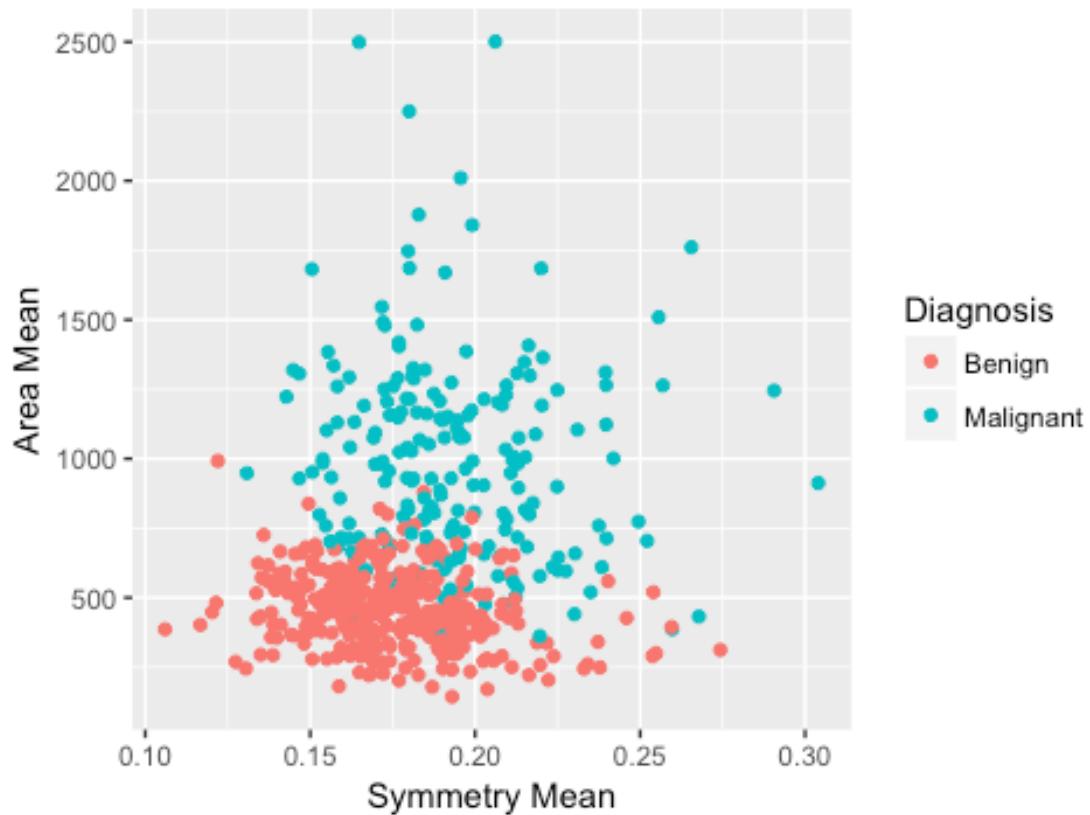
```
#perimeter_mean vs. symmetry_mean
symmetry_perimeter_mean_scatter <- ggplot(data = cancer, aes(x = symmetry_mean, y = perimeter_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Mean", y = "Perimeter Mean", title = "Perimeter Mean vs. Symmetry Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_perimeter_mean_scatter
```

Perimeter Mean vs. Symmetry Mean



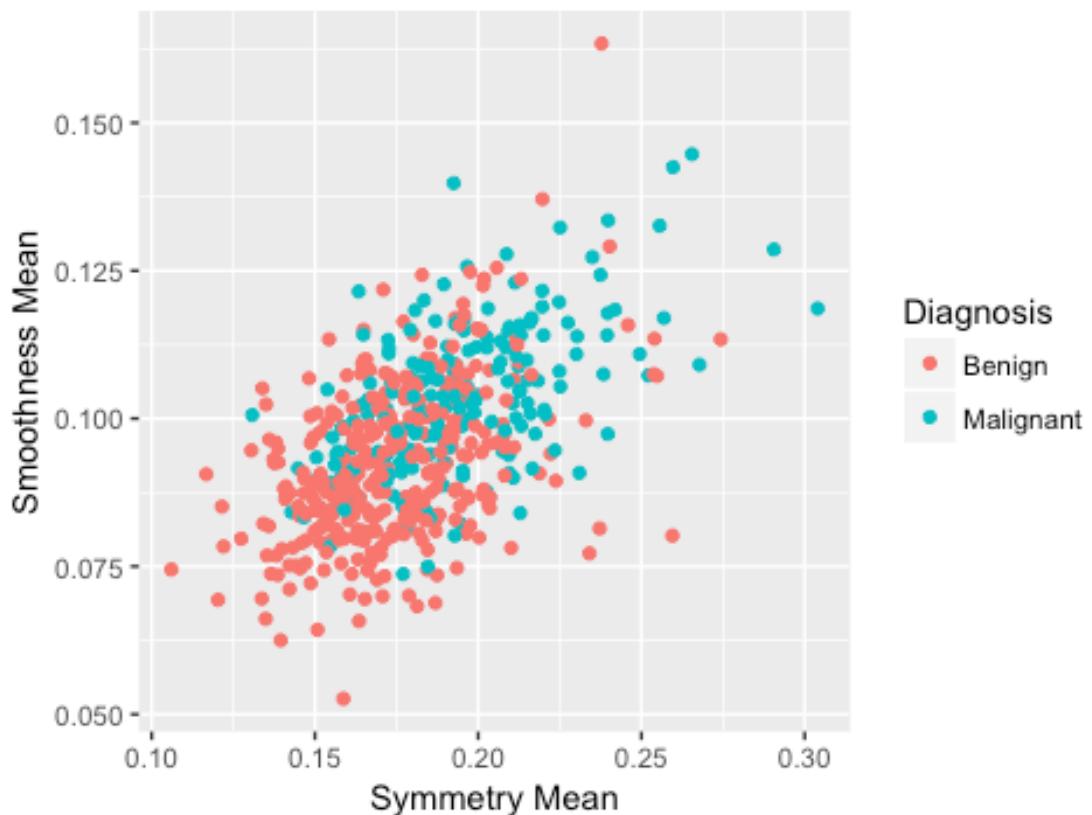
```
#area_mean vs. symmetry_mean
symmetry_area_mean_scatter <- ggplot(data = cancer, aes(x = symmetry_mean, y = area_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Mean", y = "Area Mean", title = "Area Mean vs. Symmetry Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_area_mean_scatter
```

Area Mean vs. Symmetry Mean



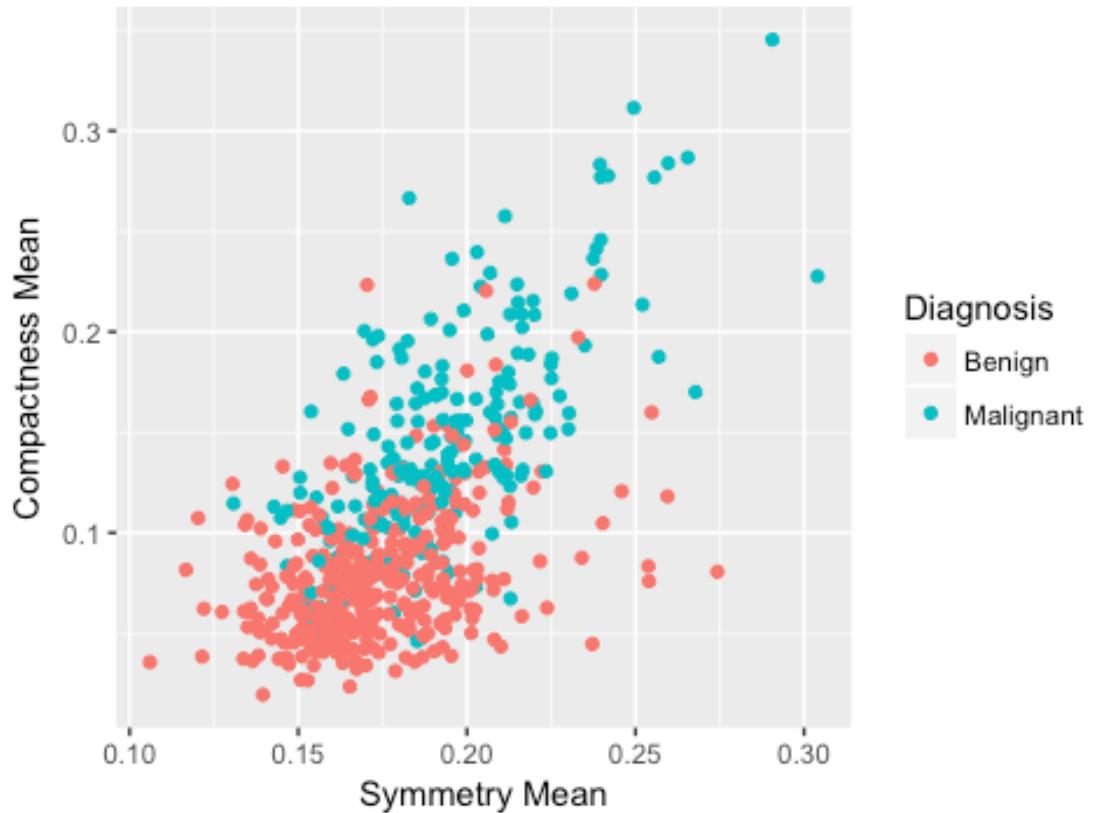
```
#smoothness_mean vs. symmetry_mean
symmetry_smoothness_mean_scatter <- ggplot(data = cancer, aes(x = symmetry_mean, y = smoothness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Mean", y = "Smoothness Mean", title = "Smoothness Mean vs. Symmetry Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_smoothness_mean_scatter
```

Smoothness Mean vs. Symmetry Mean



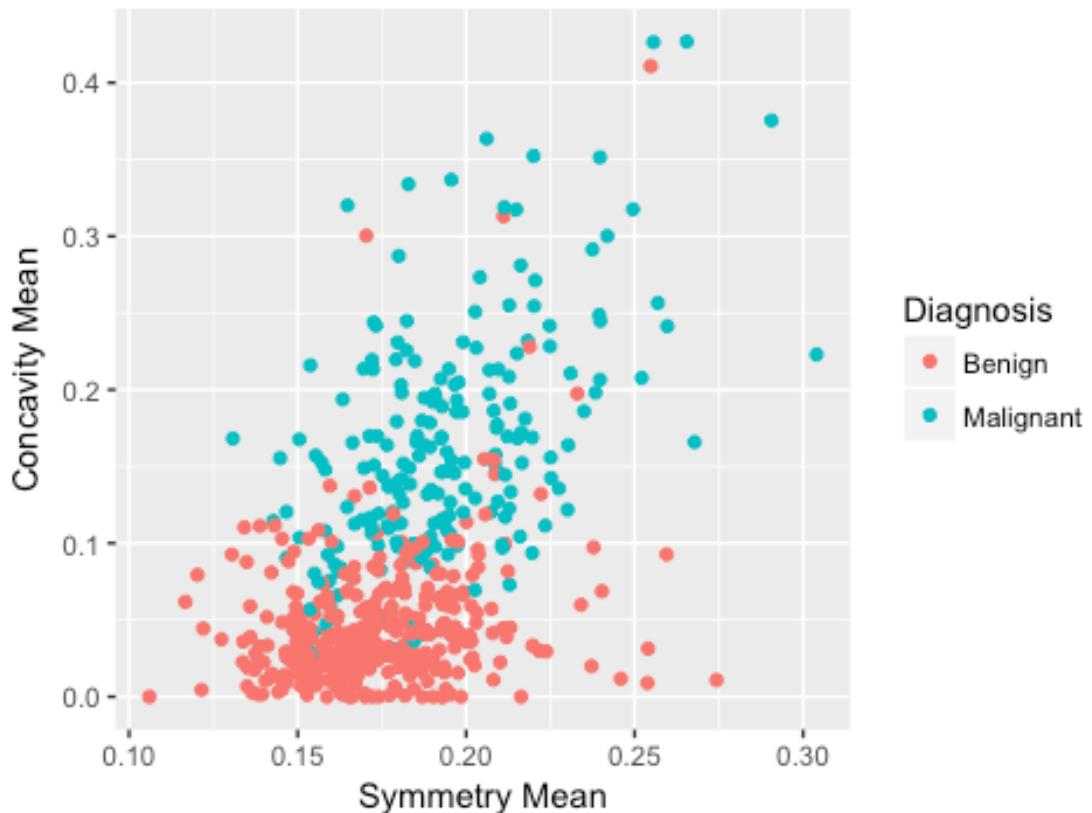
```
#compactness_mean vs. symmetry_mean
symmetry_compactness_mean_scatter <- ggplot(data = cancer, aes(x = symmetry_mean, y = compactness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Mean", y = "Compactness Mean", title = "Compactness Mean vs. Symmetry Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_compactness_mean_scatter
```

Compactness Mean vs. Symmetry Mean



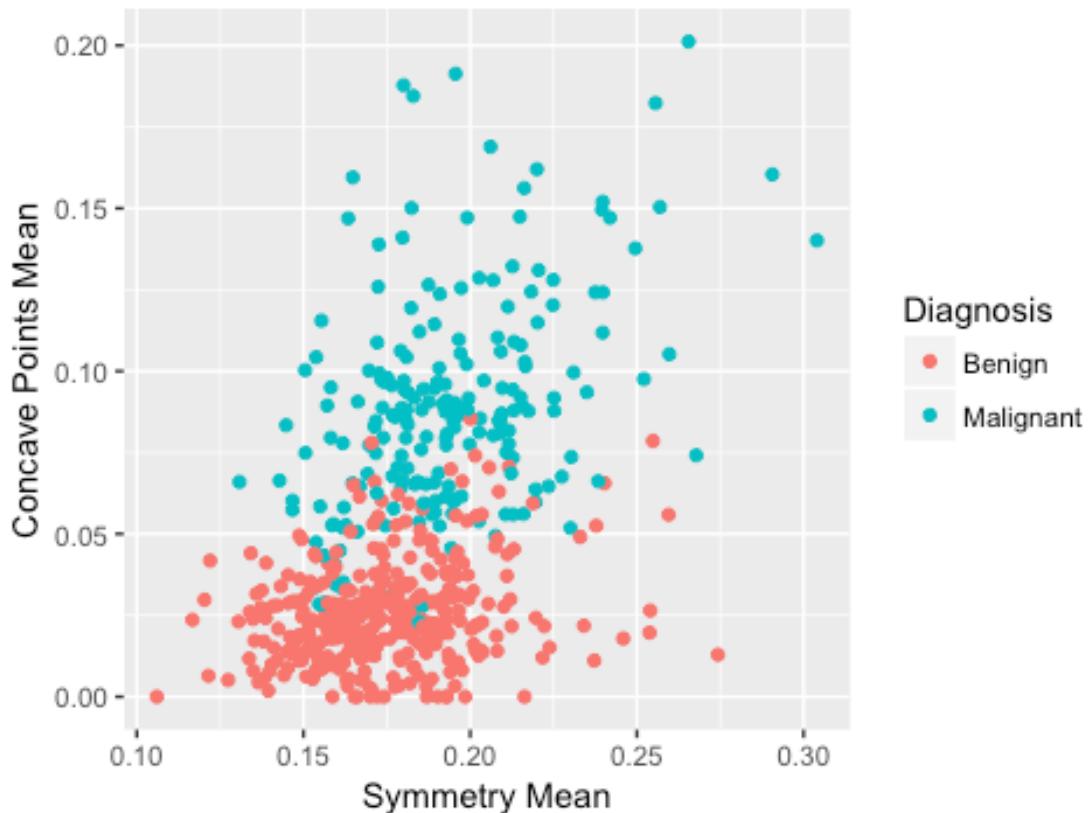
```
#concavity_mean vs. symmetry_mean
symmetry_concavity_mean_scatter <- ggplot(data = cancer, aes(x = symmetry_mean, y = concavity_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Mean", y = "Concavity Mean", title = "Concavity Mean vs. Symmetry Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_concavity_mean_scatter
```

Concavity Mean vs. Symmetry Mean



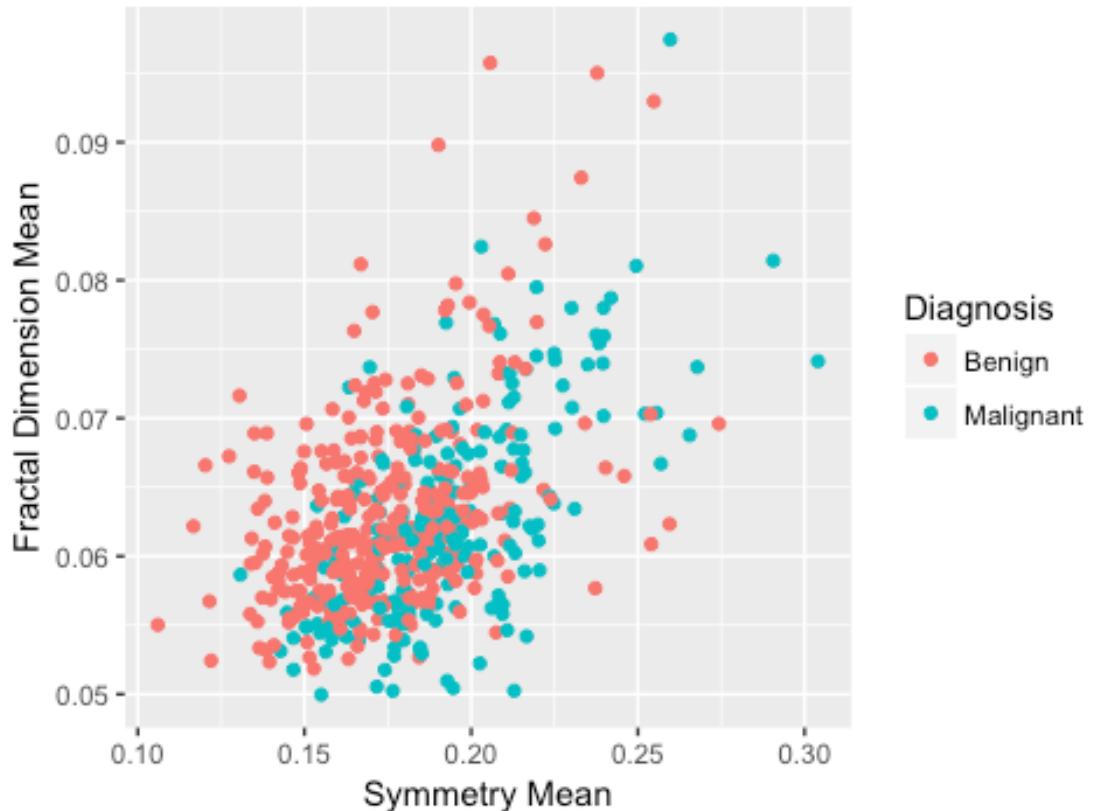
```
#concave.points_mean vs. symmetry_mean
symmetry_concave.points_mean_scatter <- ggplot(data = cancer, aes(x = symmetry_mean, y = concave.points_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Mean", y = "Concave Points Mean", title =
  "Concave Points Mean vs. Symmetry Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_concave.points_mean_scatter
```

Concave Points Mean vs. Symmetry Mean



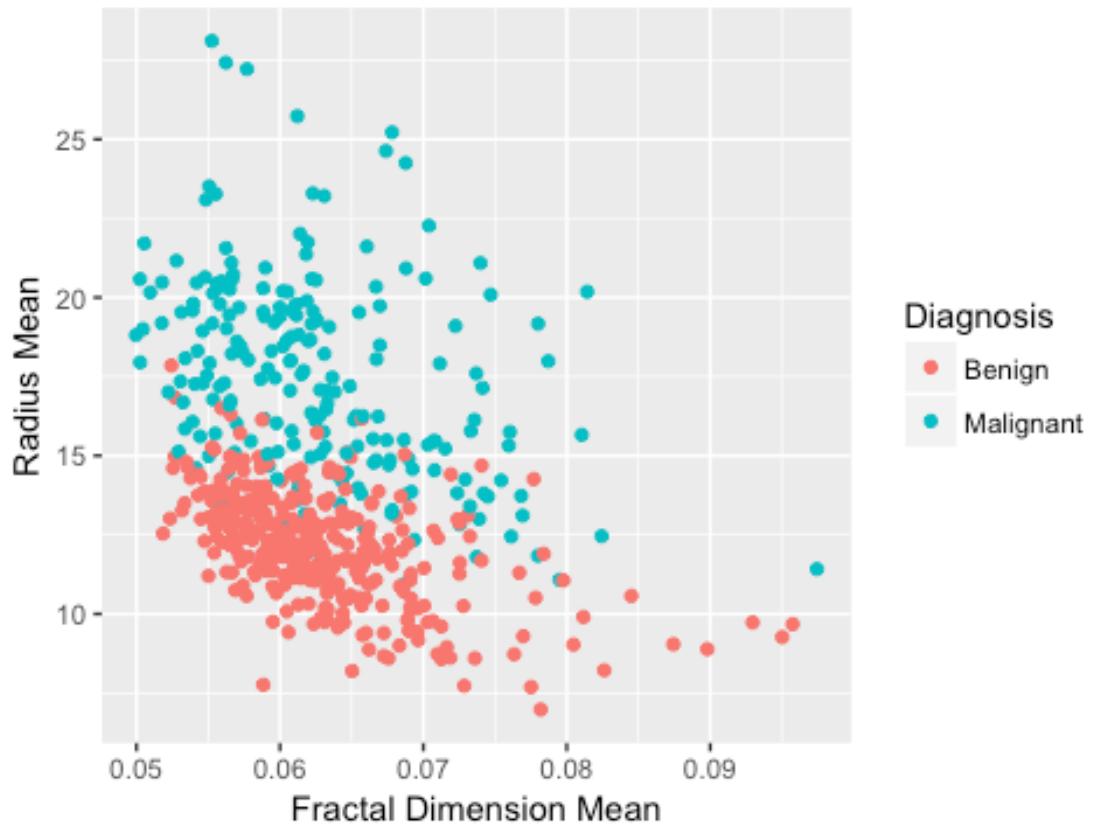
```
#fractal_dimension_mean vs. symmetry_mean
symmetry_fractal_dimension_mean_scatter <- ggplot(data = cancer, aes(x = symmetry_mean, y = fractal_dimension_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Mean", y = "Fractal Dimension Mean", title = "Fractal Dimension Mean vs. Symmetry Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_fractal_dimension_mean_scatter
```

Fractal Dimension Mean vs. Symmetry Mean



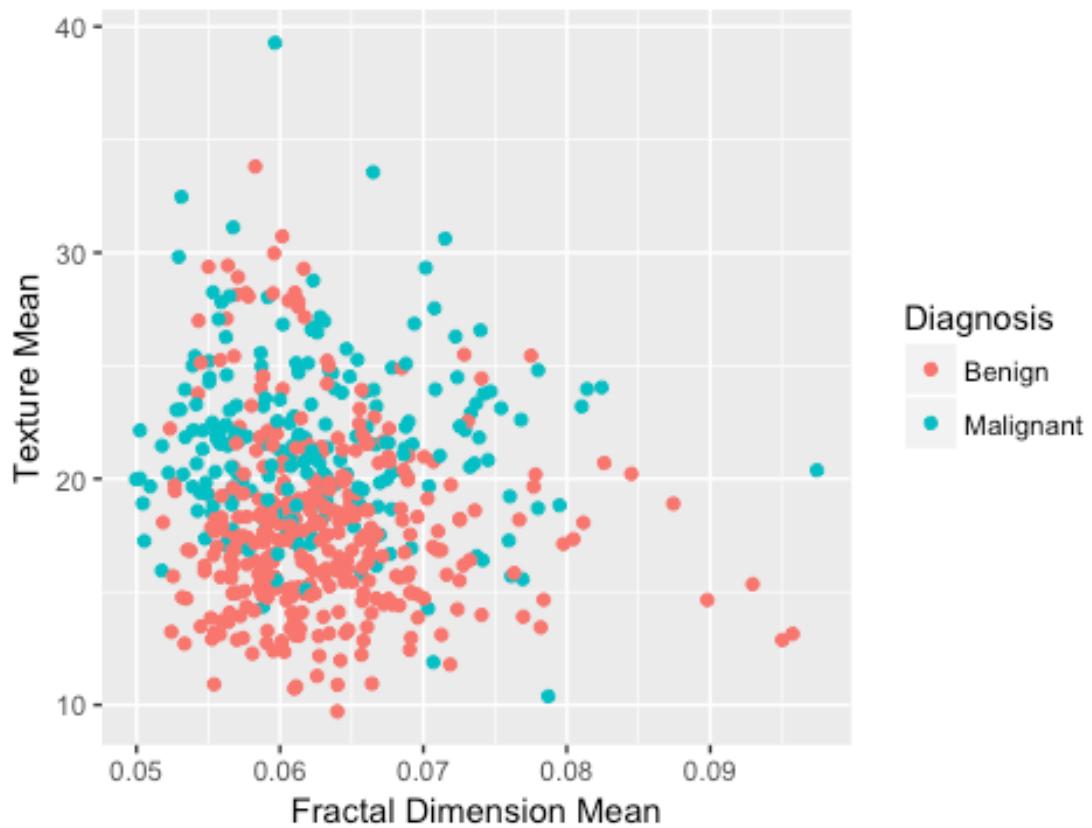
```
#radius_mean vs. fractal_dimension_mean
fractal_dimension_radius_mean_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_mean, y = radius_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Mean", y = "Radius Mean", title =
  "Radius Mean vs. Fractal Dimension Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_radius_mean_scatter
```

Radius Mean vs. Fractal Dimension Mean



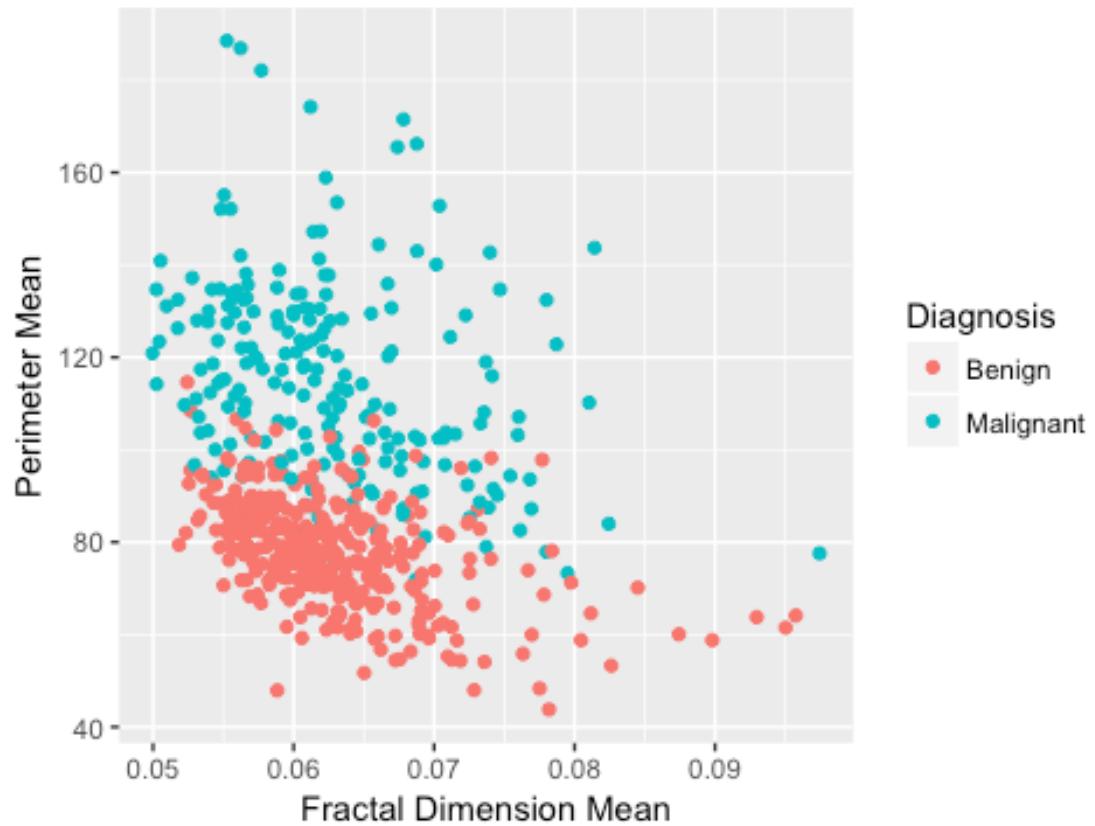
```
#texture_mean vs. fractal_dimension_mean
fractal_dimension_texture_mean_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_mean, y = texture_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Mean", y = "Texture Mean", title =
  "Texture Mean vs. Fractal Dimension Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_texture_mean_scatter
```

Texture Mean vs. Fractal Dimension Mean



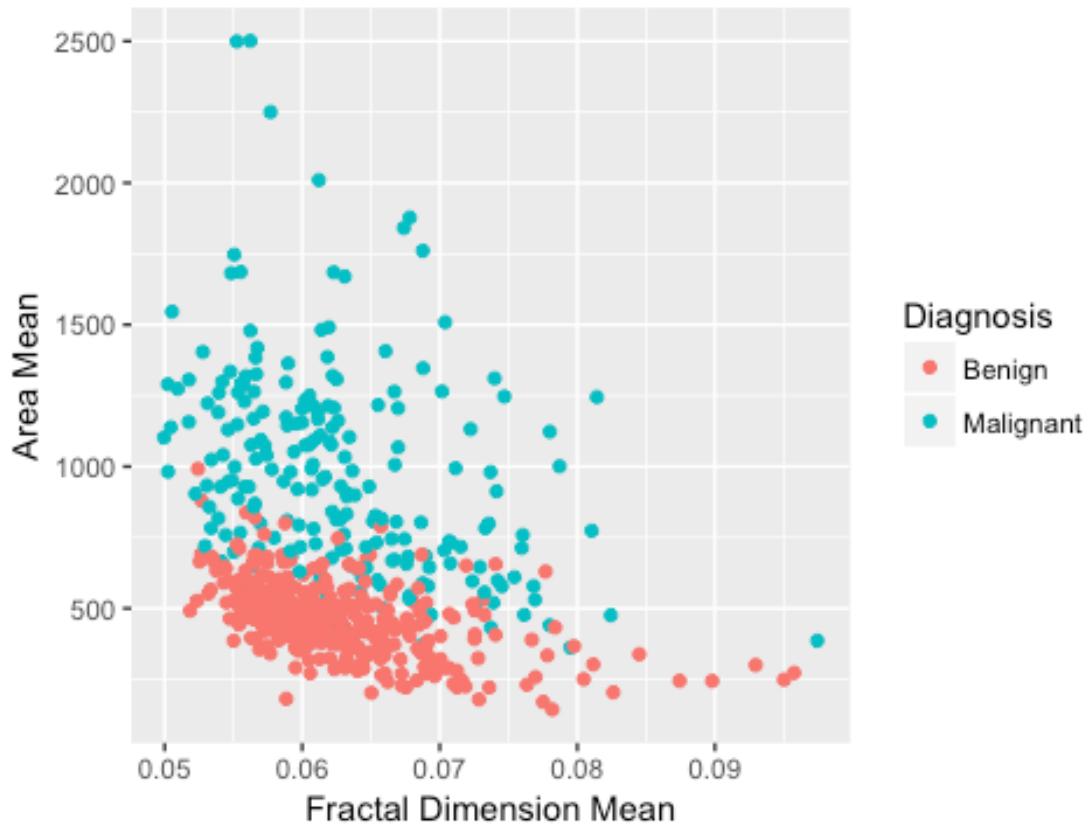
```
#perimeter_mean vs. fractal_dimension_mean
fractal_dimension_perimeter_mean_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_mean, y = perimeter_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Mean", y = "Perimeter Mean", title = "Perimeter Mean vs. Fractal Dimension Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_perimeter_mean_scatter
```

Perimeter Mean vs. Fractal Dimension Mean



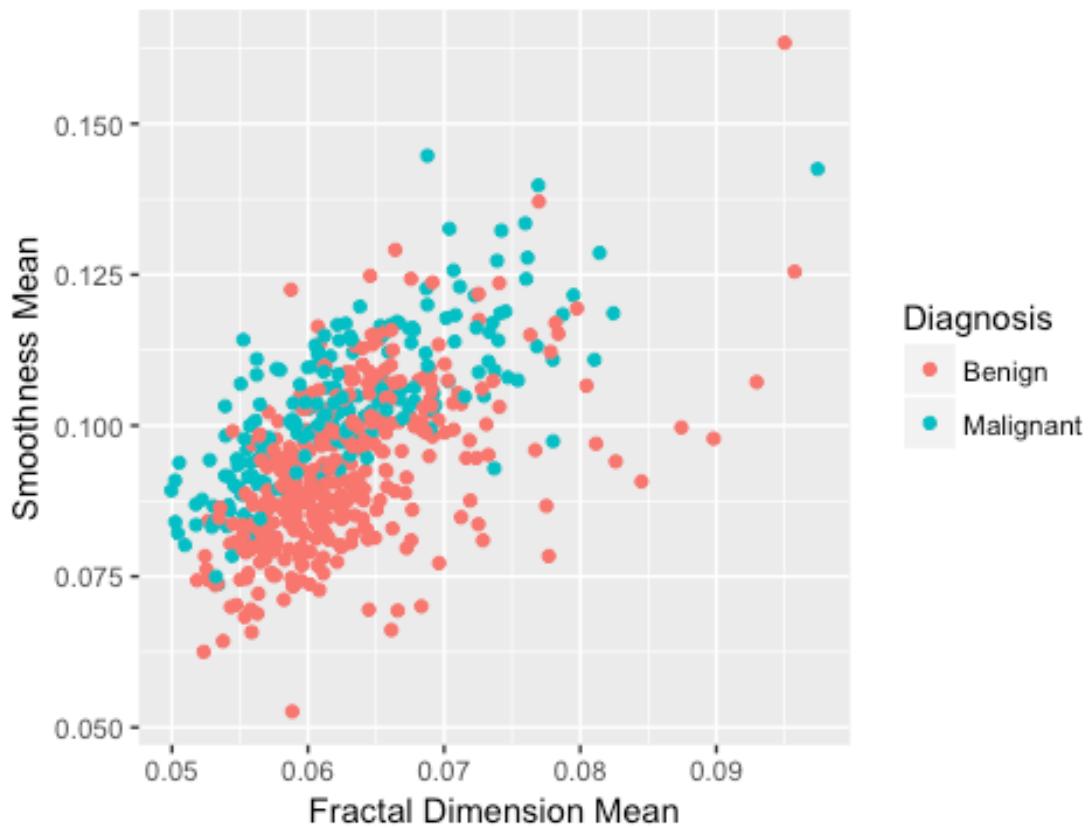
```
#area_mean vs. fractal_dimension_mean
fractal_dimension_area_mean_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_mean, y = area_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Mean", y = "Area Mean", title =
  "Area Mean vs. Fractal Dimension Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_area_mean_scatter
```

Area Mean vs. Fractal Dimension Mean



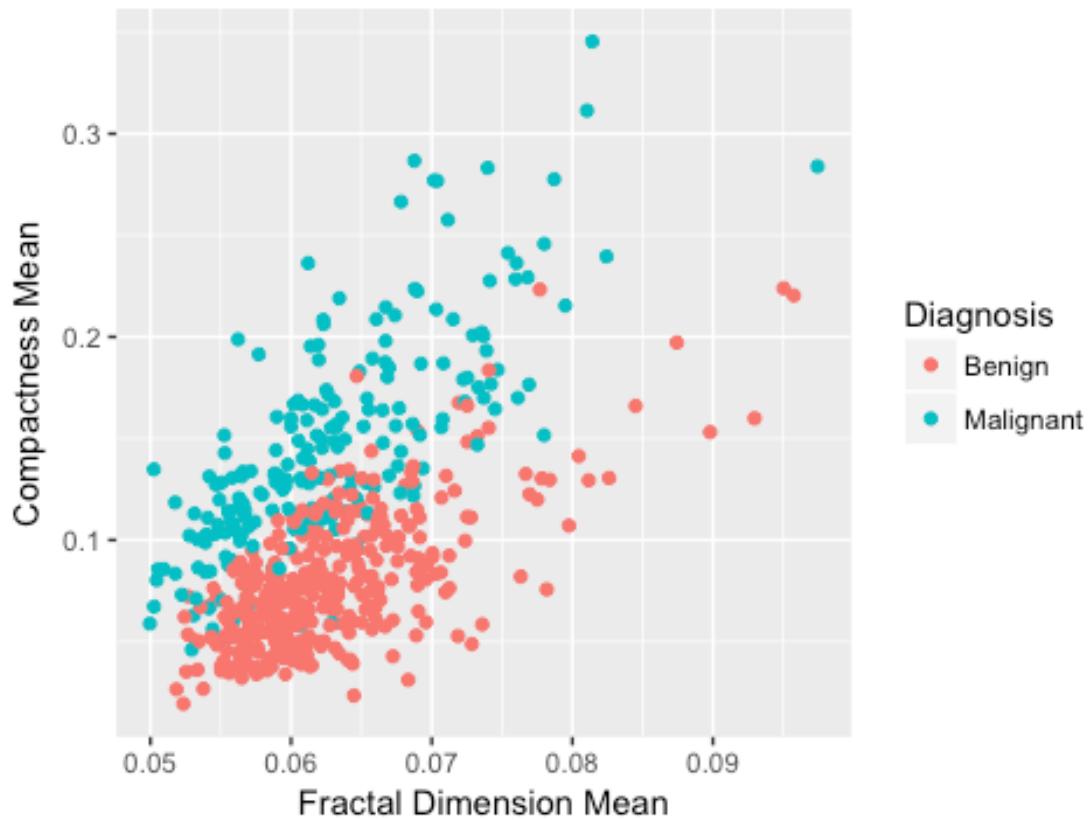
```
#smoothness_mean vs. fractal_dimension_mean
fractal_dimension_smoothness_mean_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_mean, y = smoothness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Mean", y = "Smoothness Mean", title = "Smoothness Mean vs. Fractal Dimension Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_smoothness_mean_scatter
```

Smoothness Mean vs. Fractal Dimension Mean



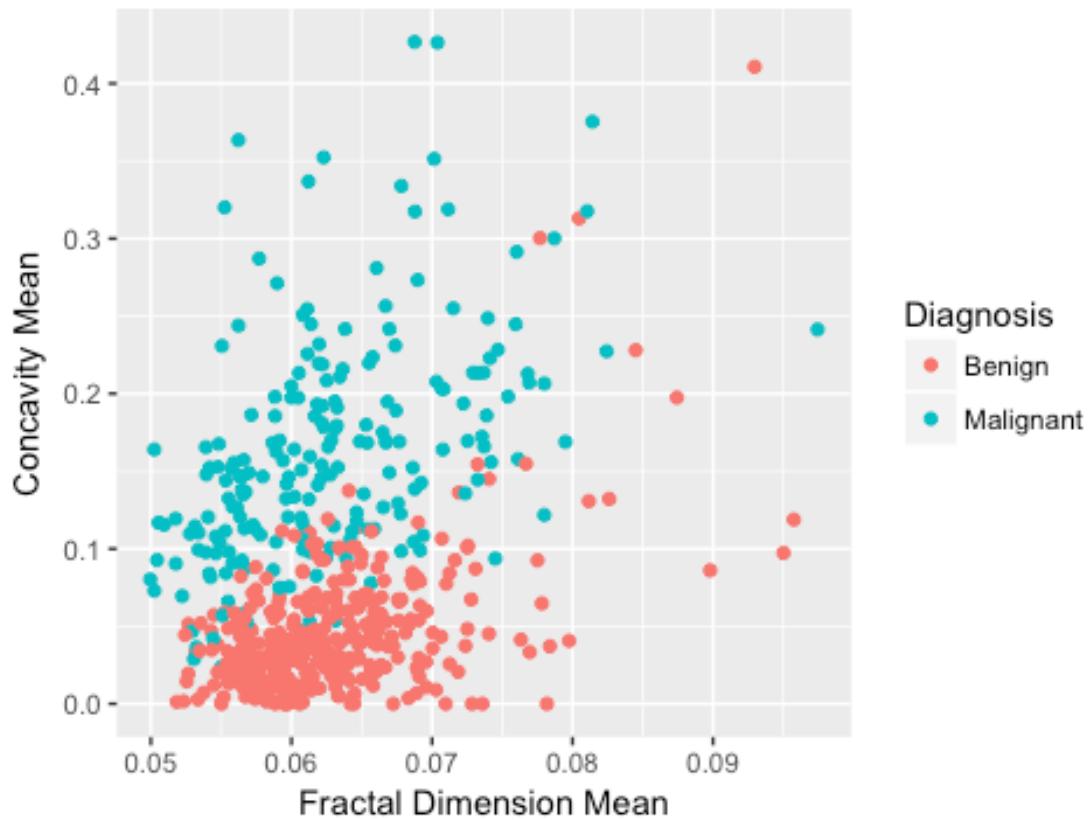
```
#compactness_mean vs. fractal_dimension_mean
fractal_dimension_compactness_mean_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_mean, y = compactness_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Mean", y = "Compactness Mean", title = "Compactness Mean vs. Fractal Dimension Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_compactness_mean_scatter
```

Compactness Mean vs. Fractal Dimension Mean



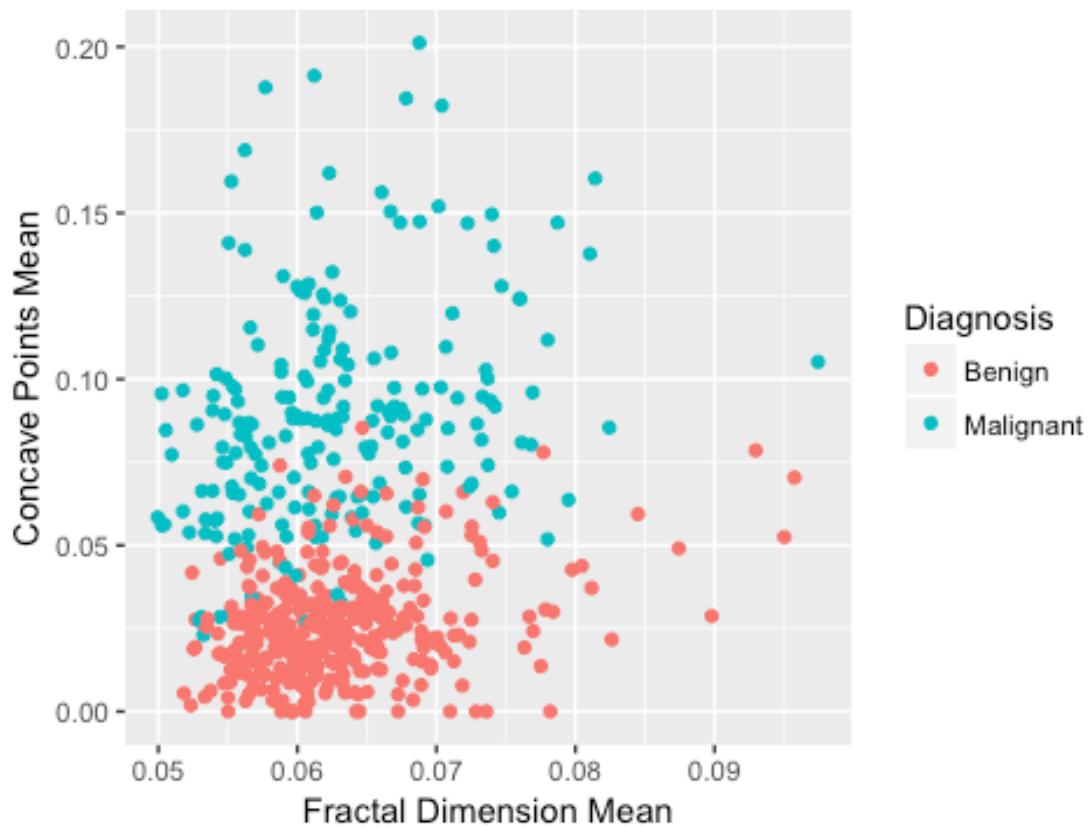
```
#concavity_mean vs. fractal_dimension_mean
fractal_dimension_concavity_mean_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_mean, y = concavity_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Mean", y = "Concavity Mean", title = "Concavity Mean vs. Fractal Dimension Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_concavity_mean_scatter
```

Concavity Mean vs. Fractal Dimension Mean



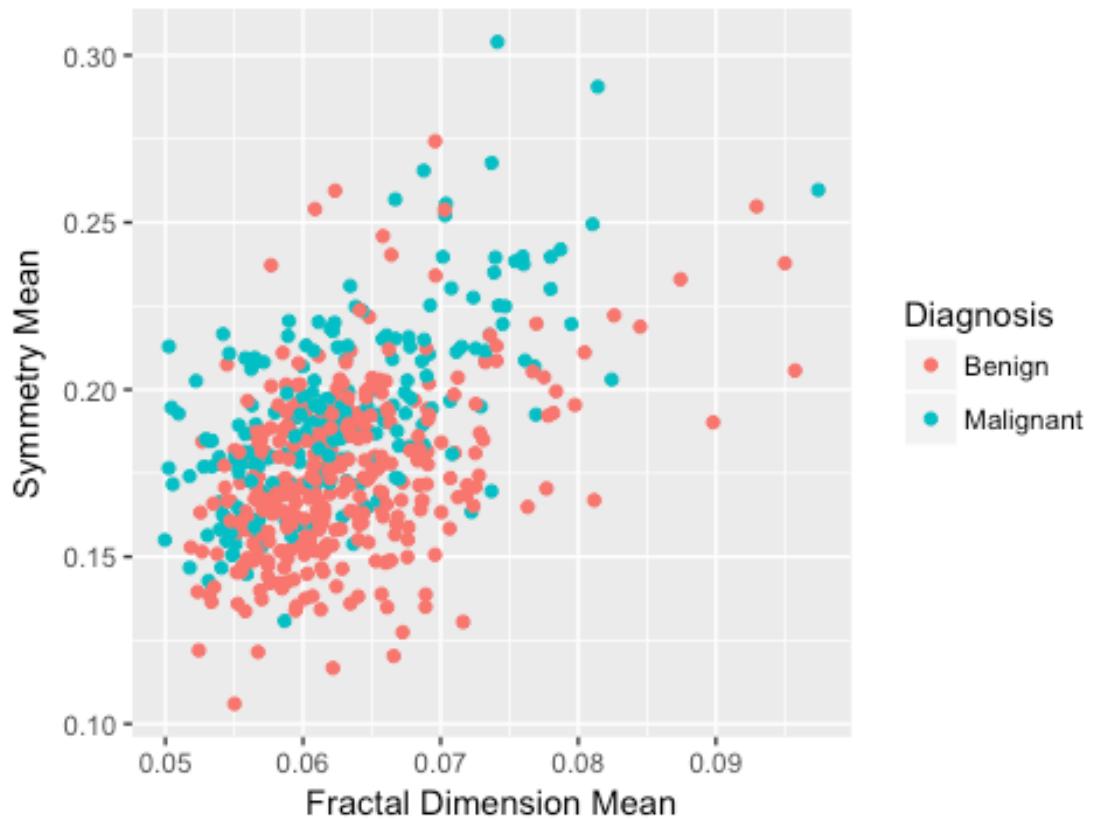
```
#concave.points_mean vs. fractal_dimension_mean
fractal_dimension_concave.points_mean_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_mean, y = concave.points_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Mean", y = "Concave Points Mean",
, title = "Concave Points Mean vs. Fractal Dimension Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_concave.points_mean_scatter
```

Concave Points Mean vs. Fractal Dimension Mean



```
#symmetry_mean vs. fractal_dimension_mean
fractal_dimension_symmetry_mean_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_mean, y = symmetry_mean, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Mean", y = "Symmetry Mean", title = "Symmetry Mean vs. Fractal Dimension Mean") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_symmetry_mean_scatter
```

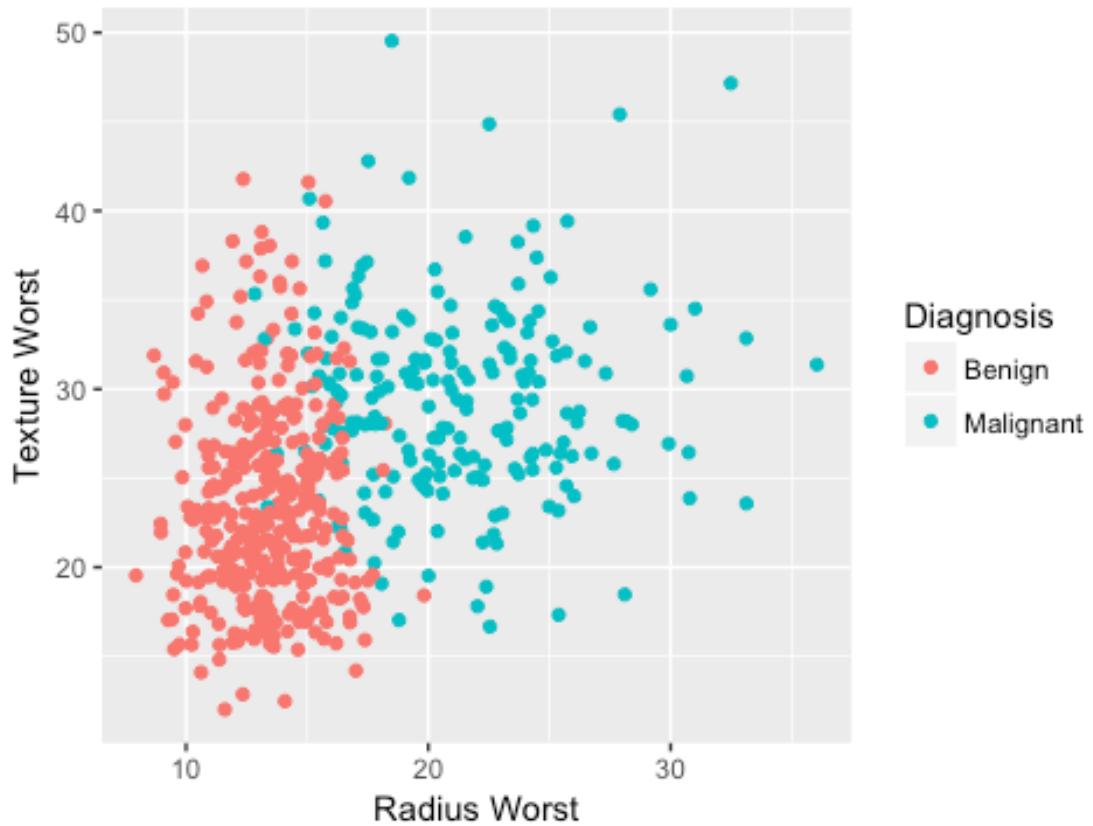
Symmetry Mean vs. Fractal Dimension Mean



```
#Worst variable relationships
```

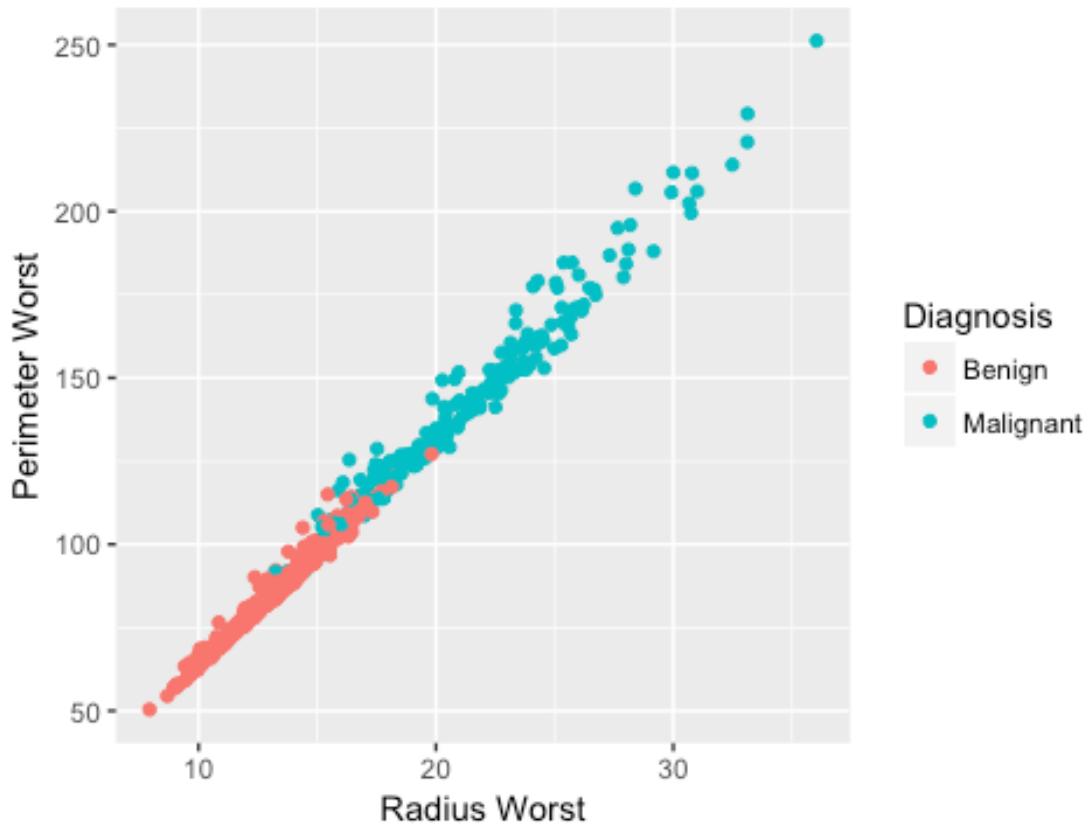
```
#texture_worst vs. radius_worst
radius_texture_worst_scatter <- ggplot(data = cancer, aes(x = radius_worst, y = texture_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Worst", y = "Texture Worst", title = "Texture Worst vs. Radius Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_texture_worst_scatter
```

Texture Worst vs. Radius Worst



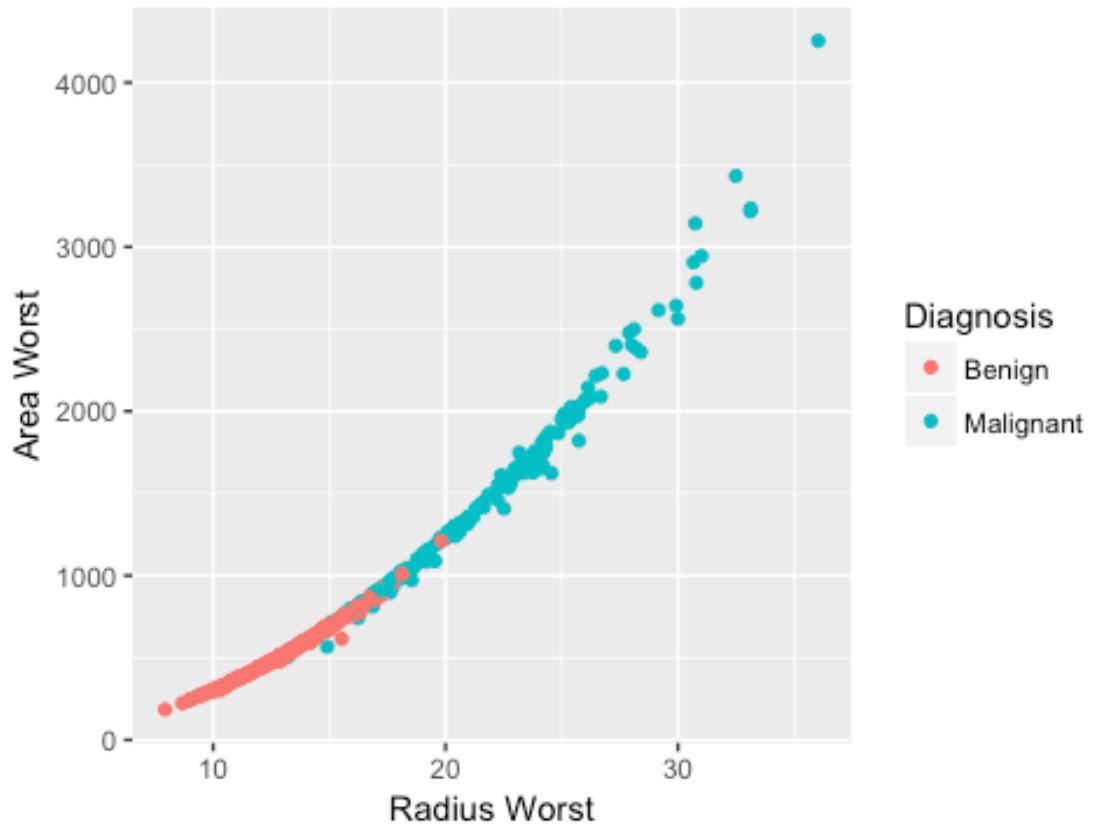
```
#perimeter_worst vs. radius_worst
radius_perimeter_worst_scatter <- ggplot(data = cancer, aes(x = radius_worst,
y = perimeter_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Worst", y = "Perimeter Worst", title = "Perimeter Worst vs. Radius Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_perimeter_worst_scatter
```

Perimeter Worst vs. Radius Worst



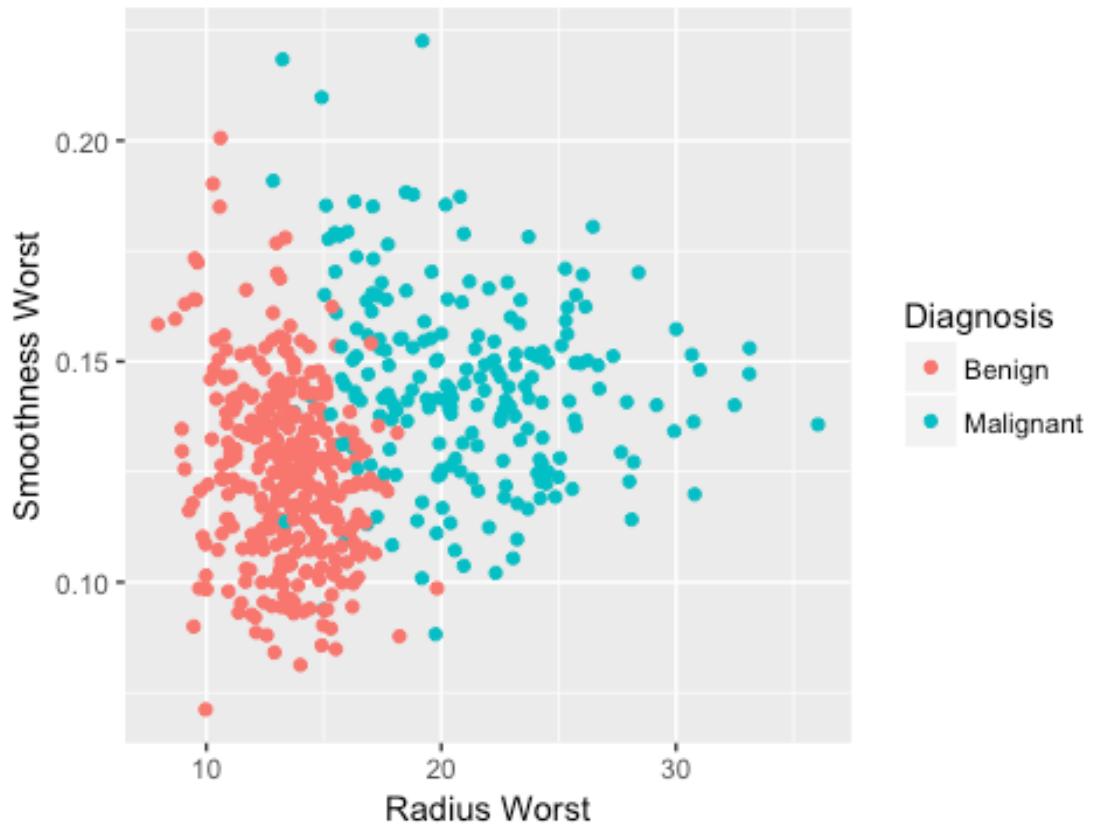
```
#area_worst vs. radius_worst
radius_area_worst_scatter <- ggplot(data = cancer, aes(x = radius_worst, y = area_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Worst", y = "Area Worst", title = "Area Worst vs. Radius Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_area_worst_scatter
```

Area Worst vs. Radius Worst



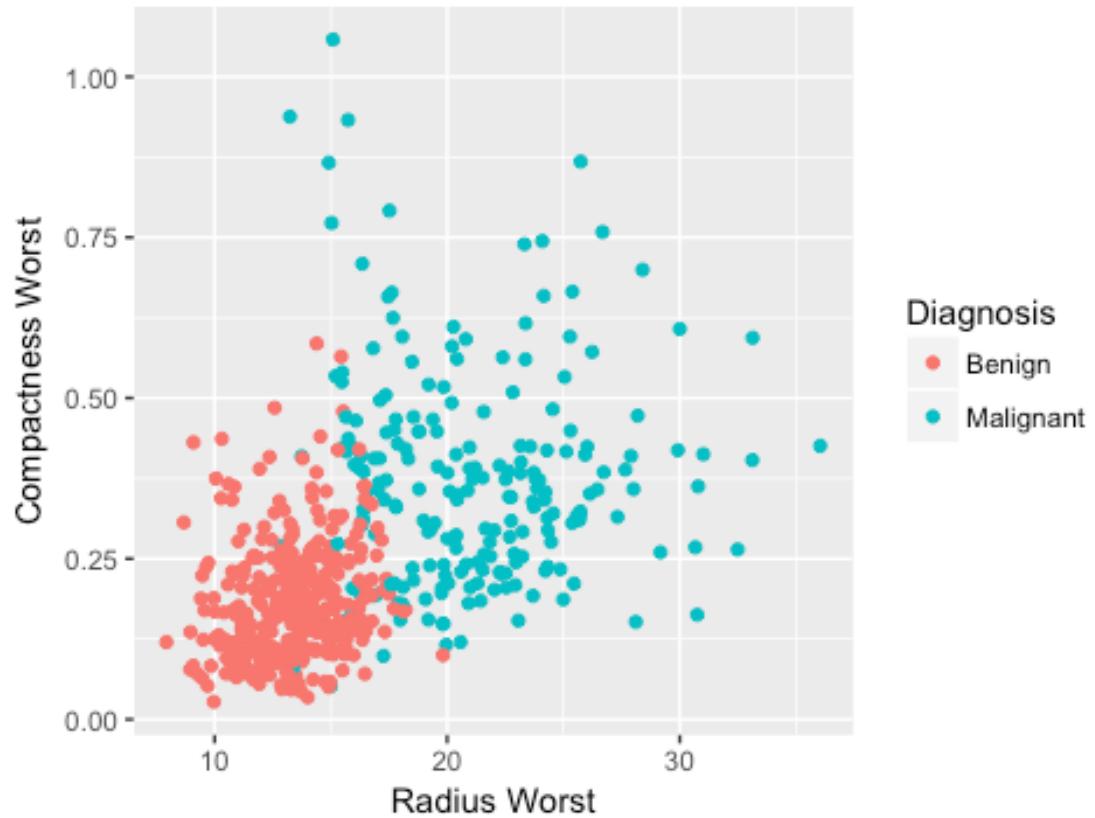
```
#smoothness_worst vs. radius_worst
radius_smoothness_worst_scatter <- ggplot(data = cancer, aes(x = radius_worst
, y = smoothness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Worst", y = "Smoothness Worst", title = "Smoothness Worst vs. Radius Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_smoothness_worst_scatter
```

Smoothness Worst vs. Radius Worst



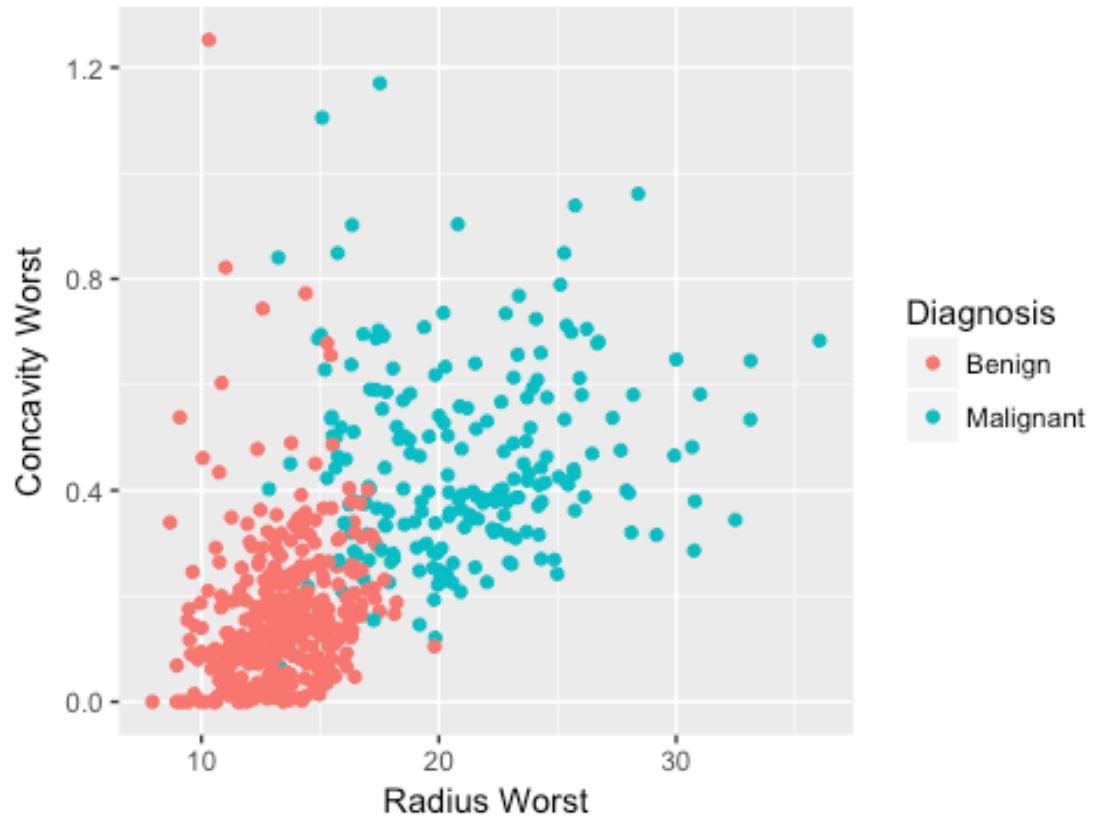
```
#compactness_worst vs. radius_worst
radius_compactness_worst_scatter <- ggplot(data = cancer, aes(x = radius_worst, y = compactness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Worst", y = "Compactness Worst", title = "Compactness Worst vs. Radius Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_compactness_worst_scatter
```

Compactness Worst vs. Radius Worst

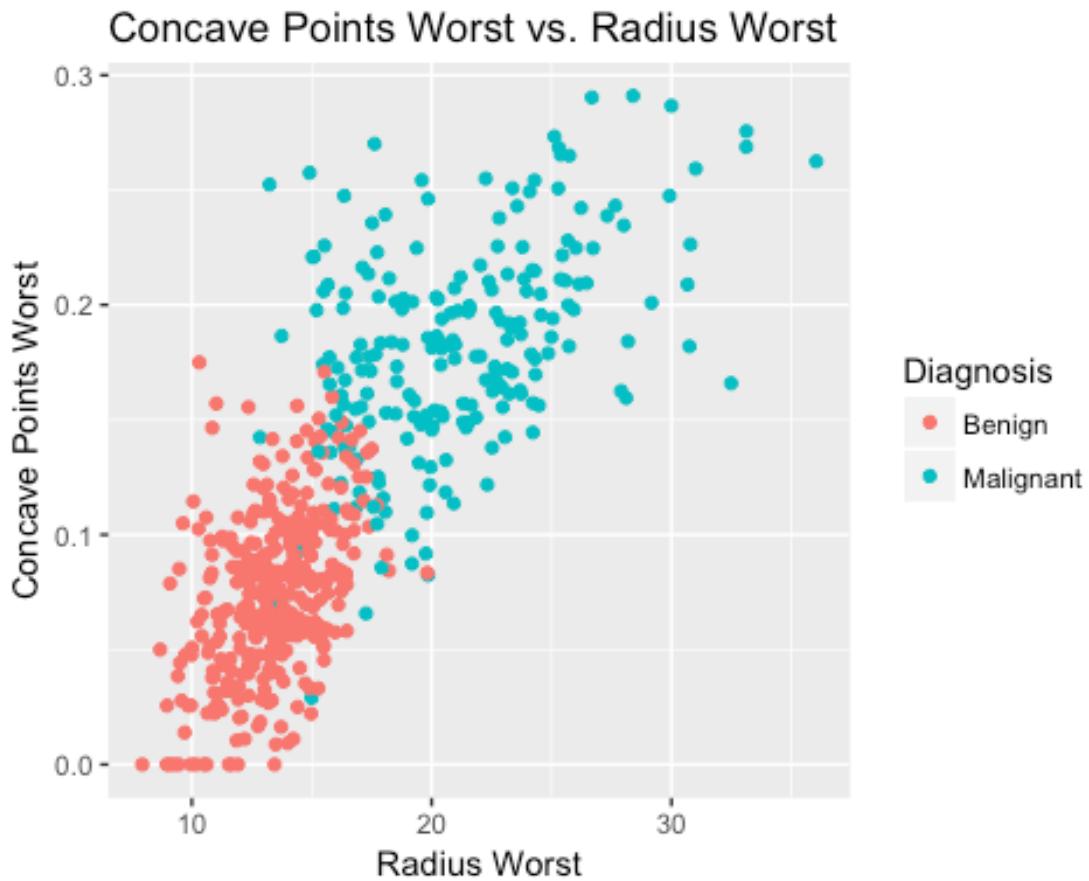


```
#concavity_worst vs. radius_worst
radius_concavity_worst_scatter <- ggplot(data = cancer, aes(x = radius_worst,
y = concavity_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Worst", y = "Concavity Worst", title = "Concavity Worst vs. Radius Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_concavity_worst_scatter
```

Concavity Worst vs. Radius Worst

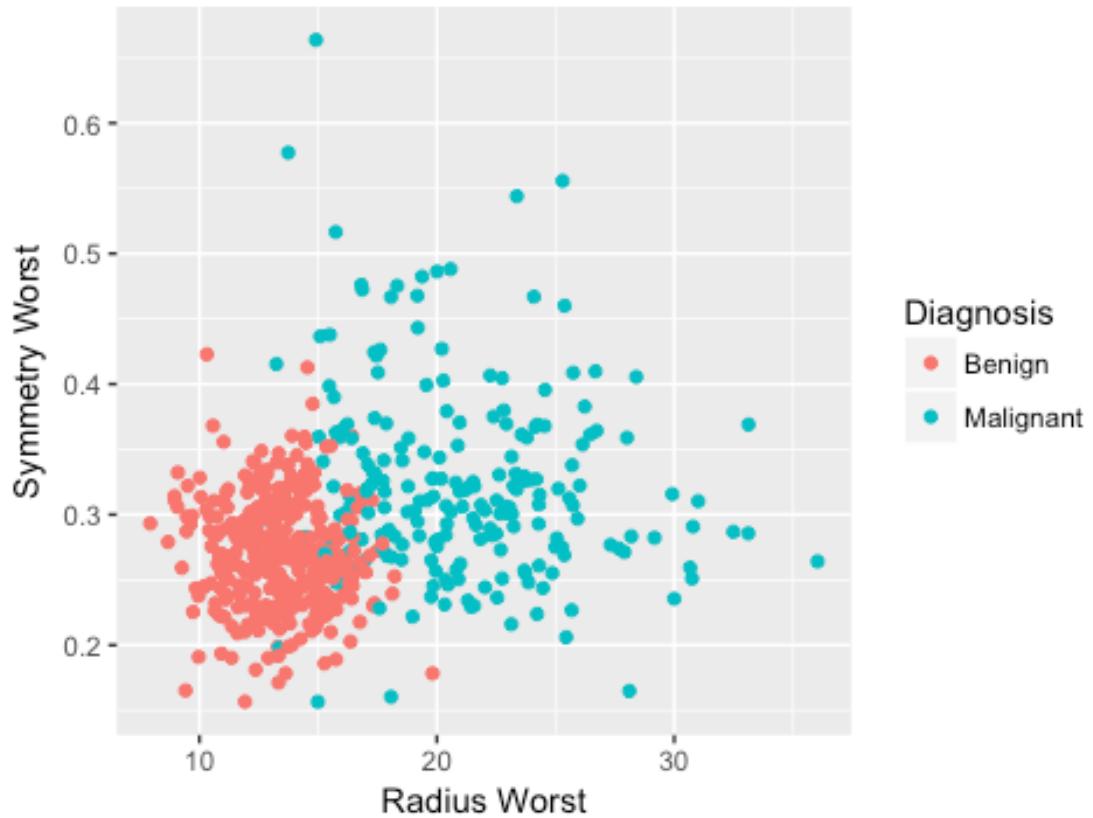


```
#concave.points_worst vs. radius_worst
radius_concave.points_worst_scatter <- ggplot(data = cancer, aes(x = radius_worst, y = concave.points_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Worst", y = "Concave Points Worst", title =
  "Concave Points Worst vs. Radius Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_concave.points_worst_scatter
```



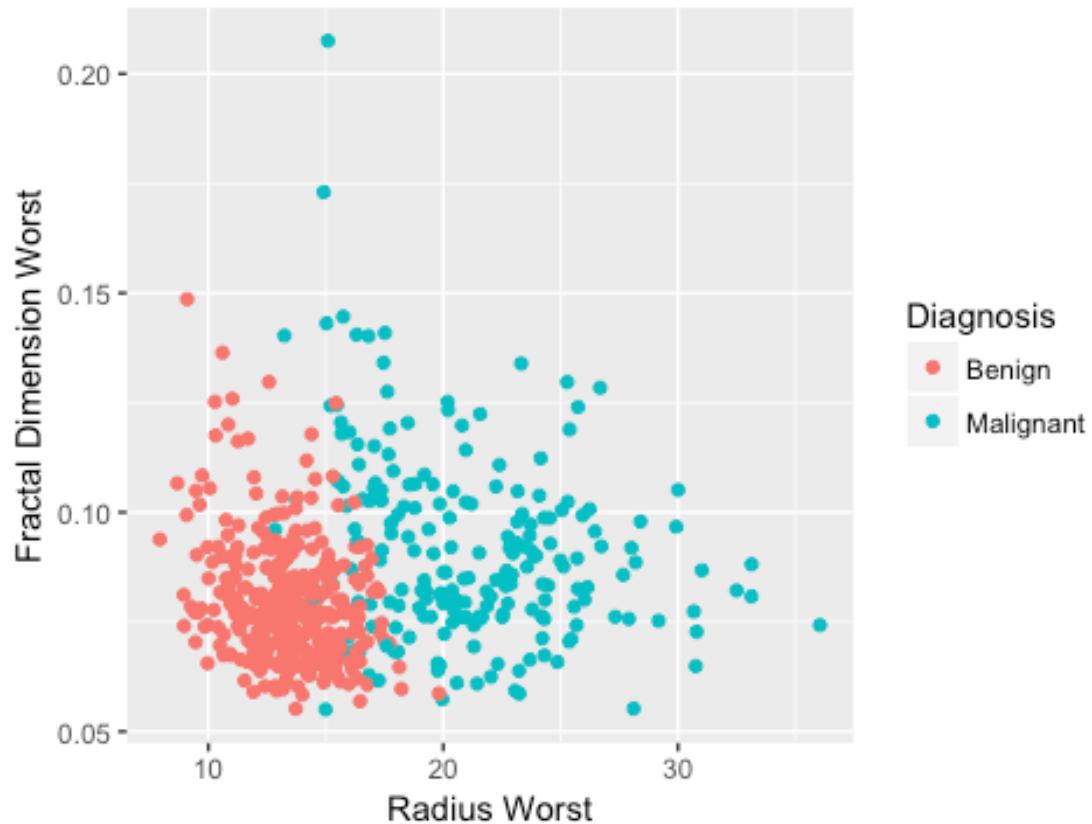
```
#symmetry_worst vs. radius_worst
radius_symmetry_worst_scatter <- ggplot(data = cancer, aes(x = radius_worst,
y = symmetry_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Worst", y = "Symmetry Worst", title = "Symmetry Worst vs. Radius Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_symmetry_worst_scatter
```

Symmetry Worst vs. Radius Worst



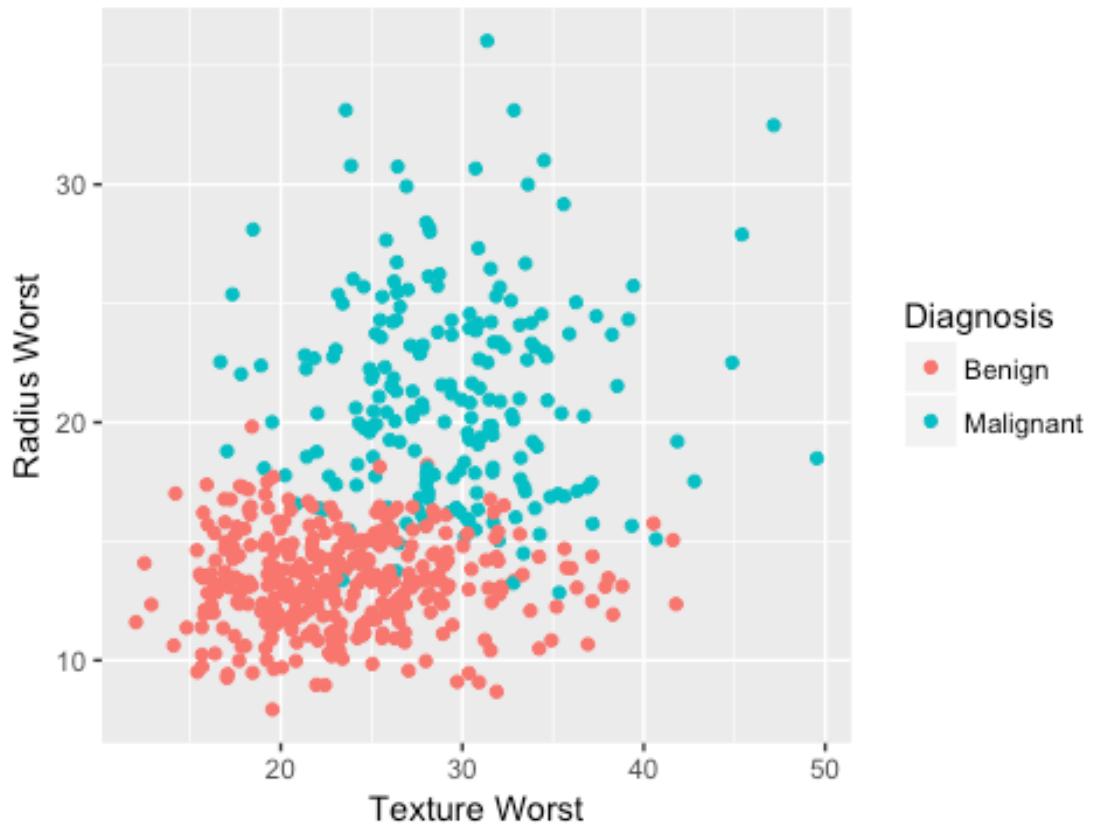
```
#fractal_dimension_worst vs. radius_worst
radius_fractal_dimension_worst_scatter <- ggplot(data = cancer, aes(x = radius_worst, y = fractal_dimension_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Worst", y = "Fractal Dimension Worst", title = "Fractal Dimension Worst vs. Radius Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_fractal_dimension_worst_scatter
```

Fractal Dimension Worst vs. Radius Worst



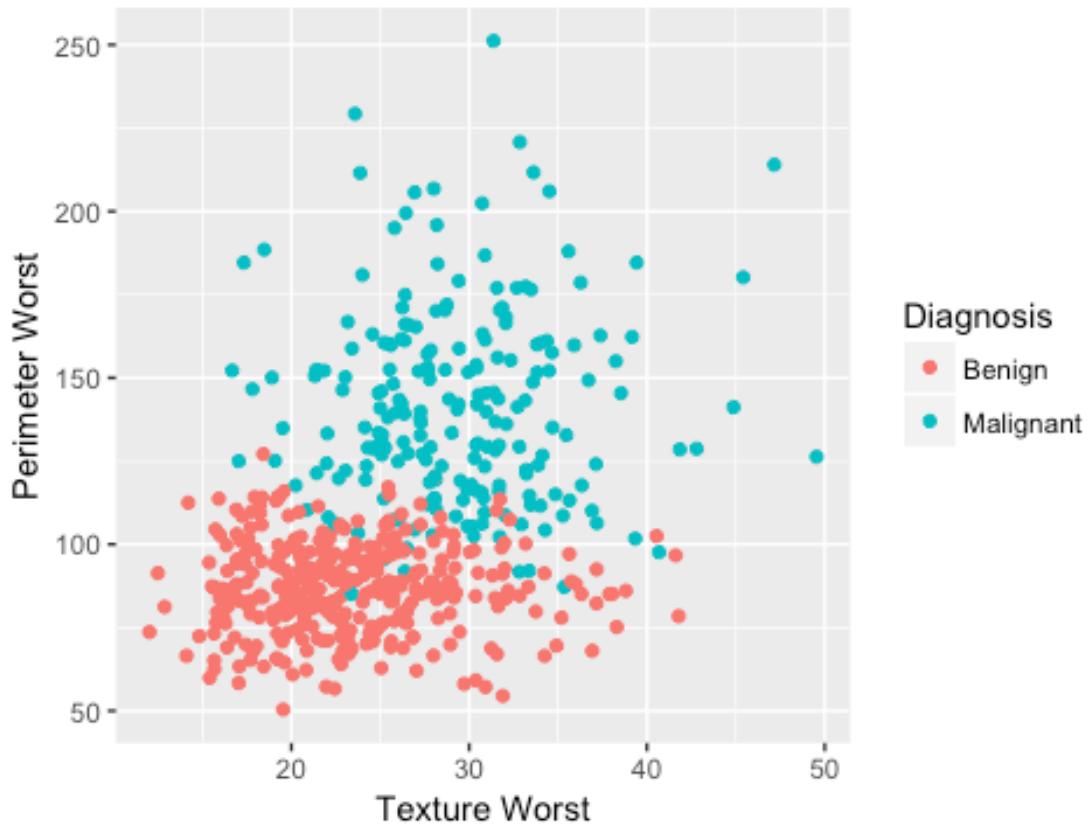
```
#radius_worst vs. texture_worst
texture_radius_worst_scatter <- ggplot(data = cancer, aes(x = texture_worst,
y = radius_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Worst", y = "Radius Worst", title = "Radius Worst vs. Texture Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_radius_worst_scatter
```

Radius Worst vs. Texture Worst



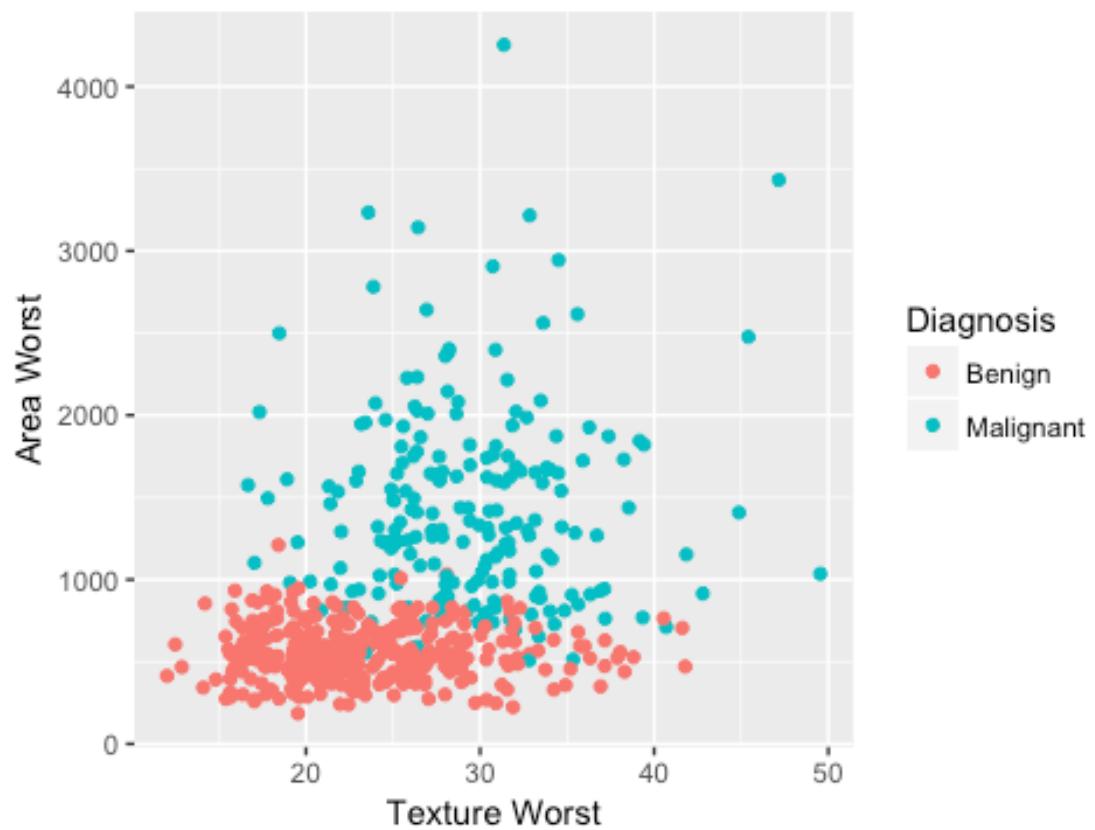
```
#perimeter_worst vs. texture_worst
texture_perimeter_worst_scatter <- ggplot(data = cancer, aes(x = texture_worst, y = perimeter_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Worst", y = "Perimeter Worst", title = "Perimeter Worst vs. Texture Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_perimeter_worst_scatter
```

Perimeter Worst vs. Texture Worst



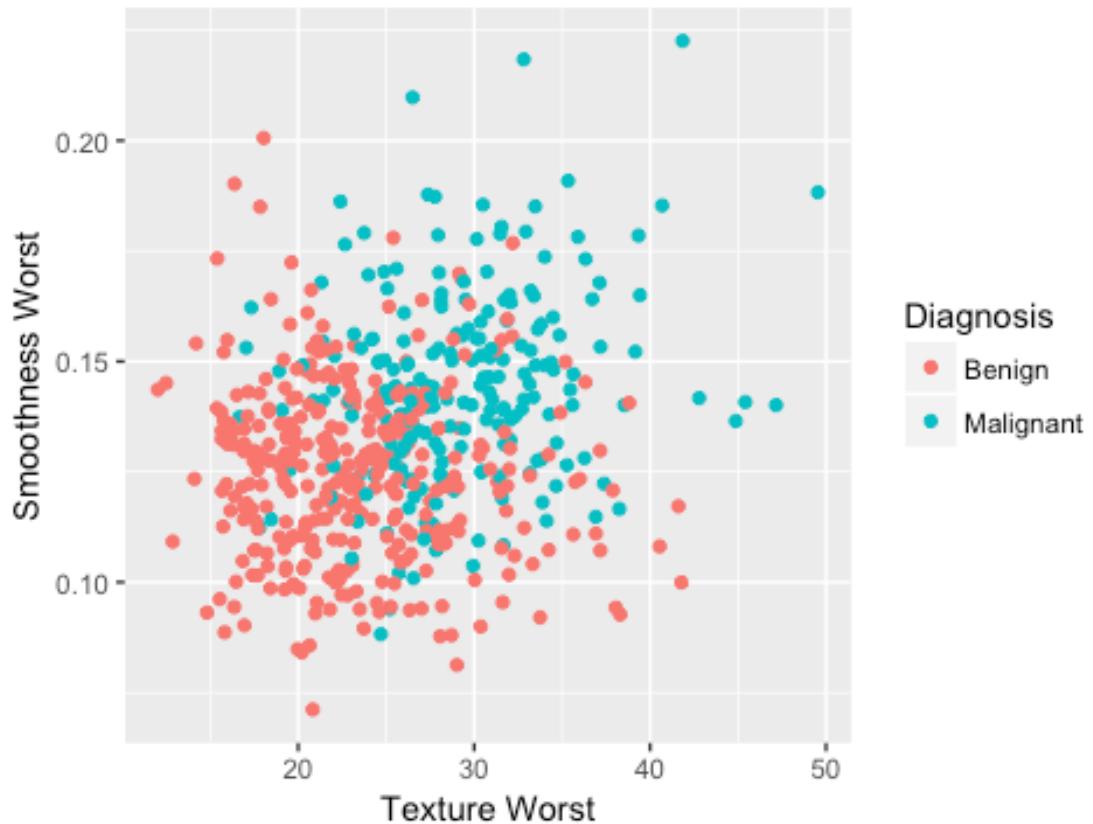
```
#area_worst vs. texture_worst
texture_area_worst_scatter <- ggplot(data = cancer, aes(x = texture_worst, y = area_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Worst", y = "Area Worst", title = "Area Worst vs. Texture Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_area_worst_scatter
```

Area Worst vs. Texture Worst



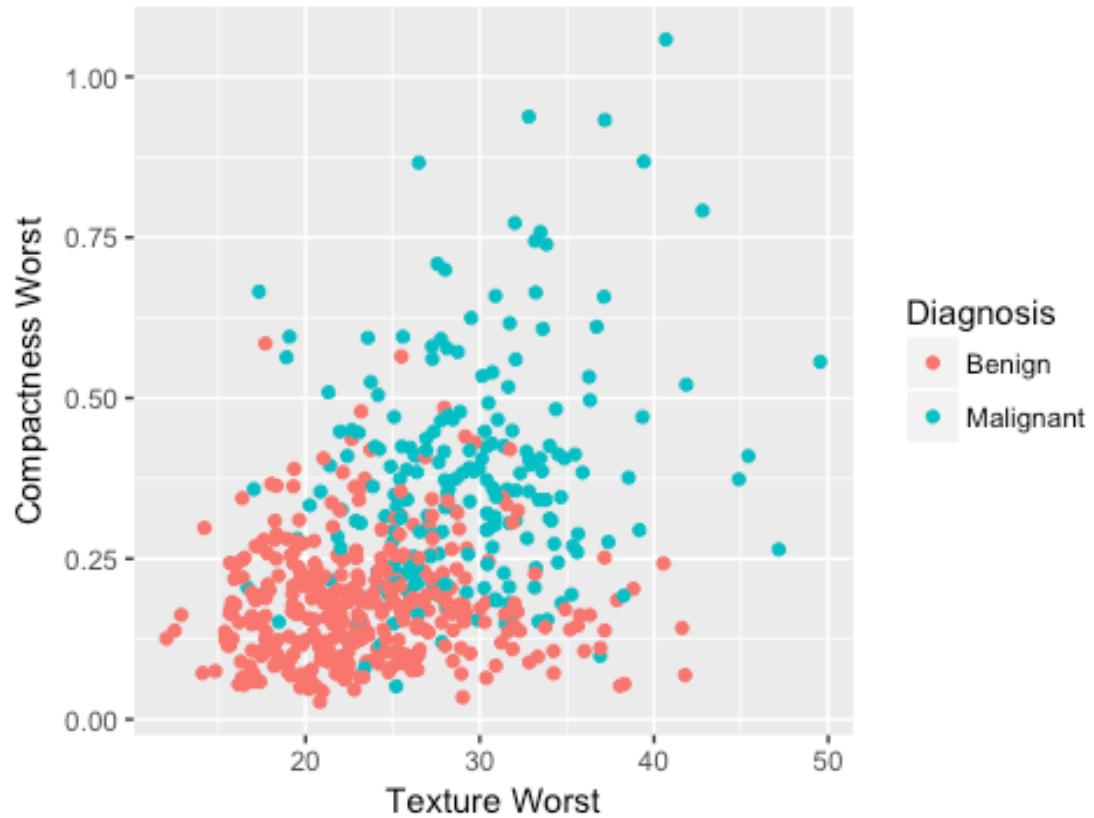
```
#smoothness_worst vs. texture_worst
texture_smoothness_worst_scatter <- ggplot(data = cancer, aes(x = texture_worst, y = smoothness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Worst", y = "Smoothness Worst", title = "Smoothness Worst vs. Texture Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_smoothness_worst_scatter
```

Smoothness Worst vs. Texture Worst



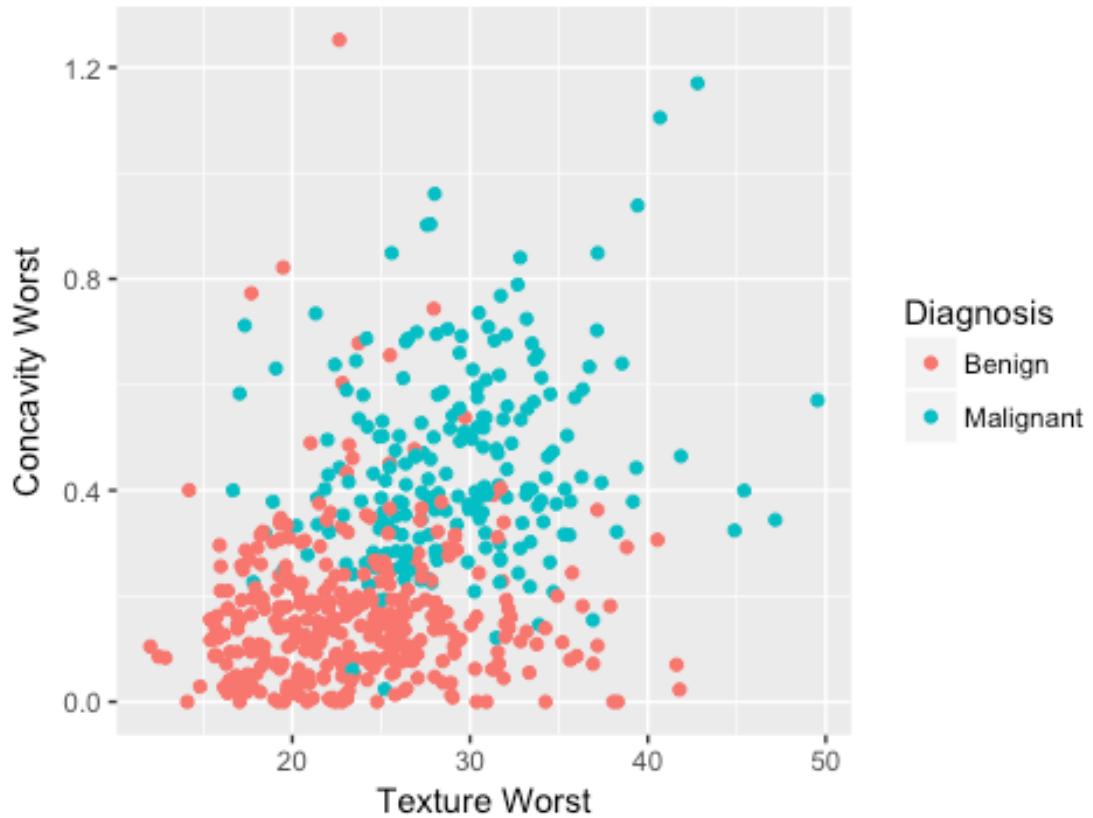
```
#compactness_worst vs. texture_worst
texture_compactness_worst_scatter <- ggplot(data = cancer, aes(x = texture_worst, y = compactness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Worst", y = "Compactness Worst", title = "Compactness Worst vs. Texture Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_compactness_worst_scatter
```

Compactness Worst vs. Texture Worst

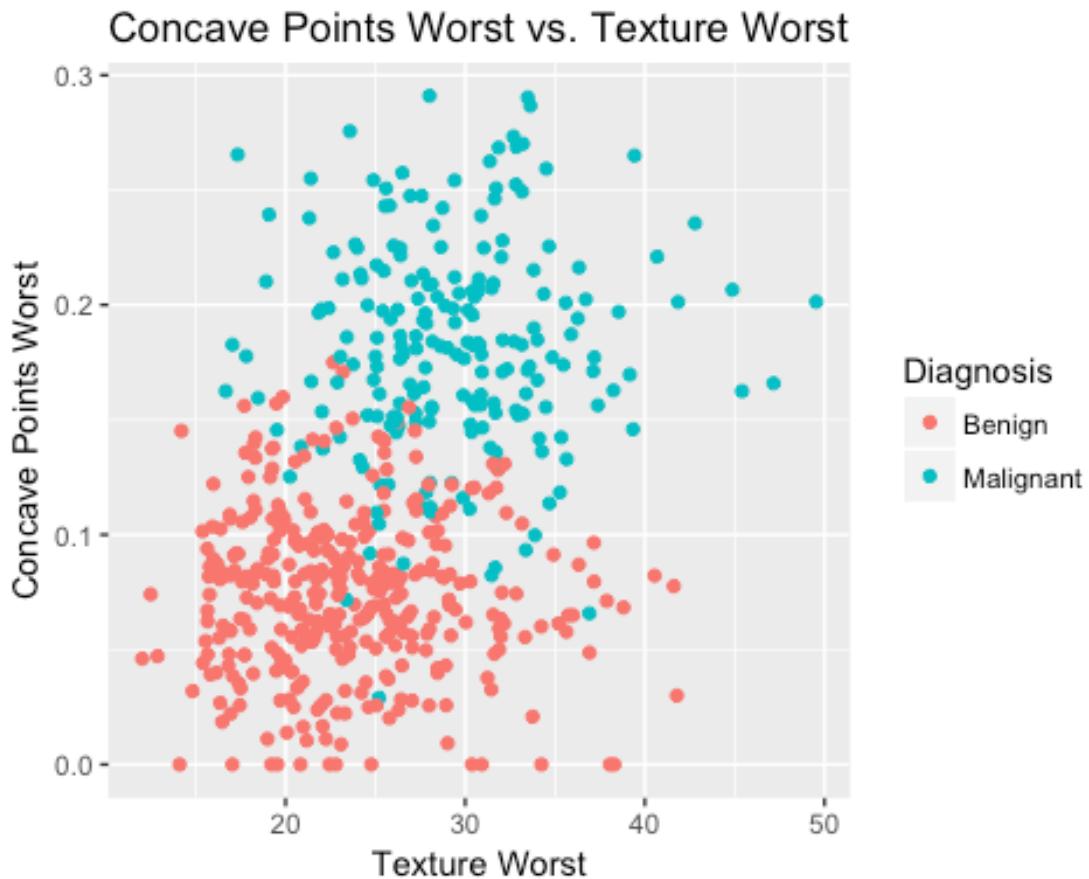


```
#concavity_worst vs. texture_worst
texture_concavity_worst_scatter <- ggplot(data = cancer, aes(x = texture_worst, y = concavity_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Worst", y = "Concavity Worst", title = "Concavity Worst vs. Texture Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_concavity_worst_scatter
```

Concavity Worst vs. Texture Worst

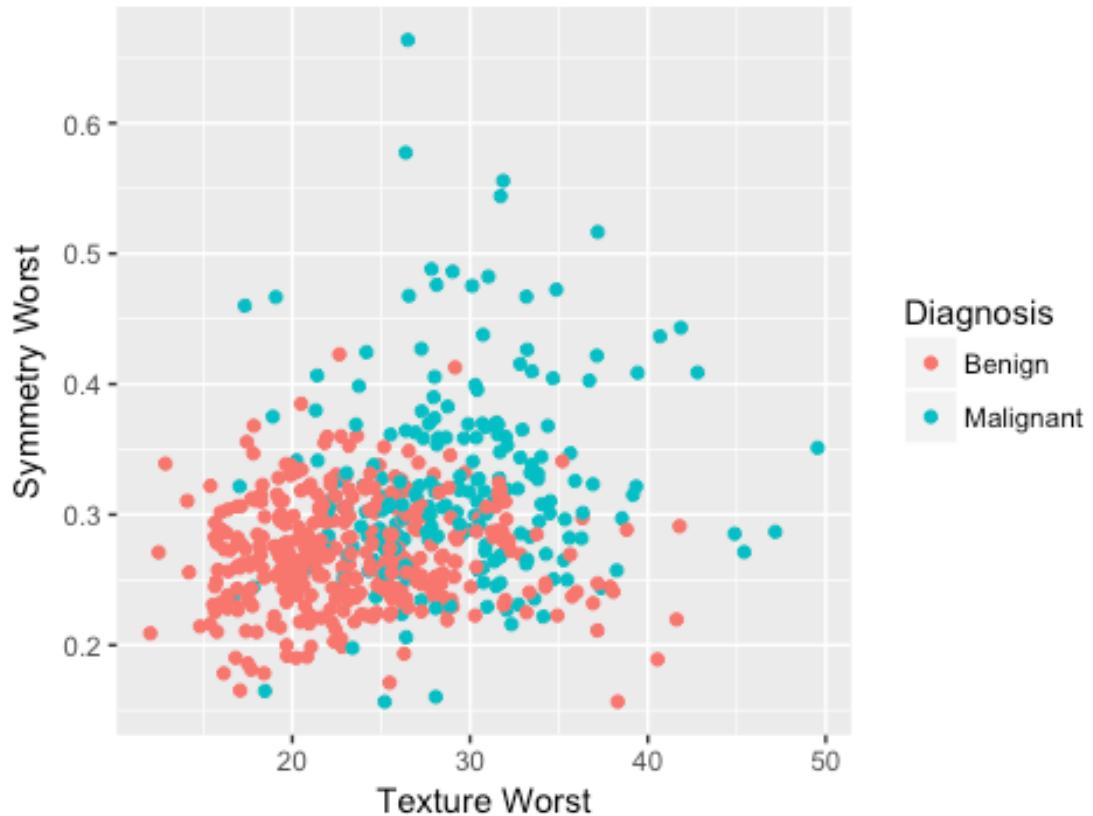


```
#concave.points_worst vs. texture_worst
texture_concave.points_worst_scatter <- ggplot(data = cancer, aes(x = texture_worst, y = concave.points_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Worst", y = "Concave Points Worst", title =
  "Concave Points Worst vs. Texture Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_concave.points_worst_scatter
```



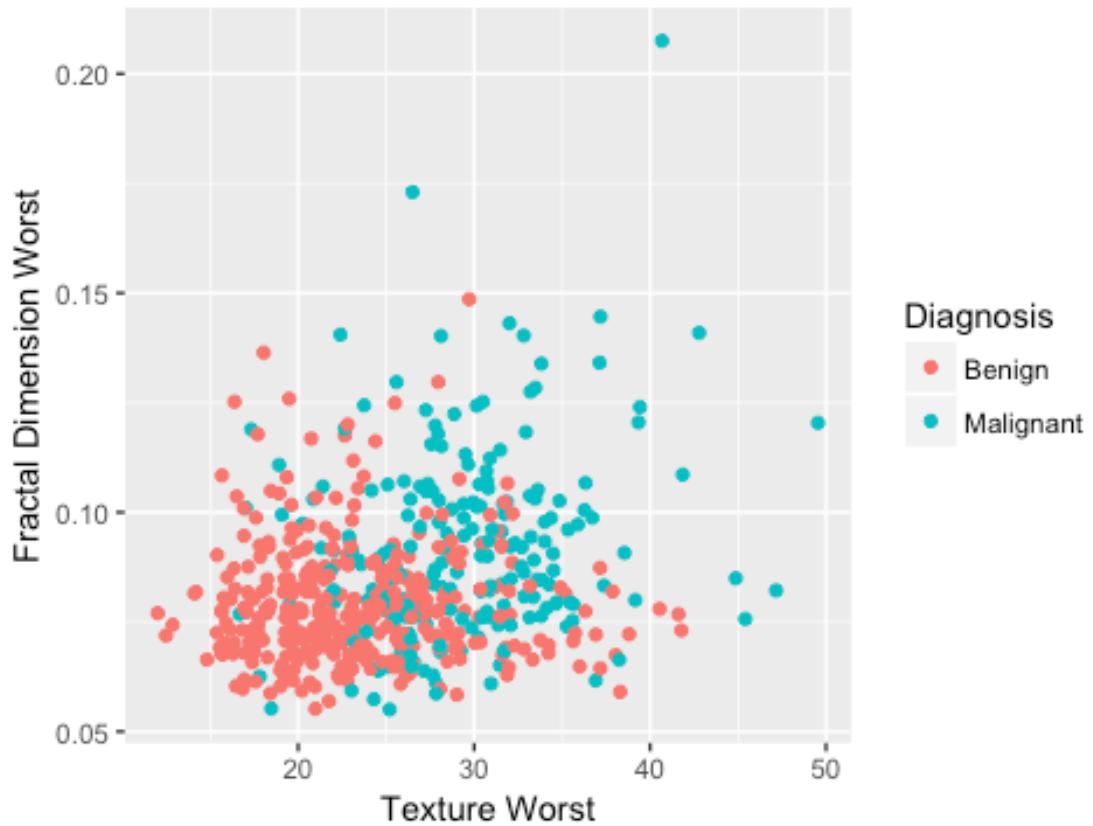
```
#symmetry_worst vs. texture_worst
texture_symmetry_worst_scatter <- ggplot(data = cancer, aes(x = texture_worst
, y = symmetry_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Worst", y = "Symmetry Worst", title = "Sym
metry Worst vs. Texture Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_symmetry_worst_scatter
```

Symmetry Worst vs. Texture Worst



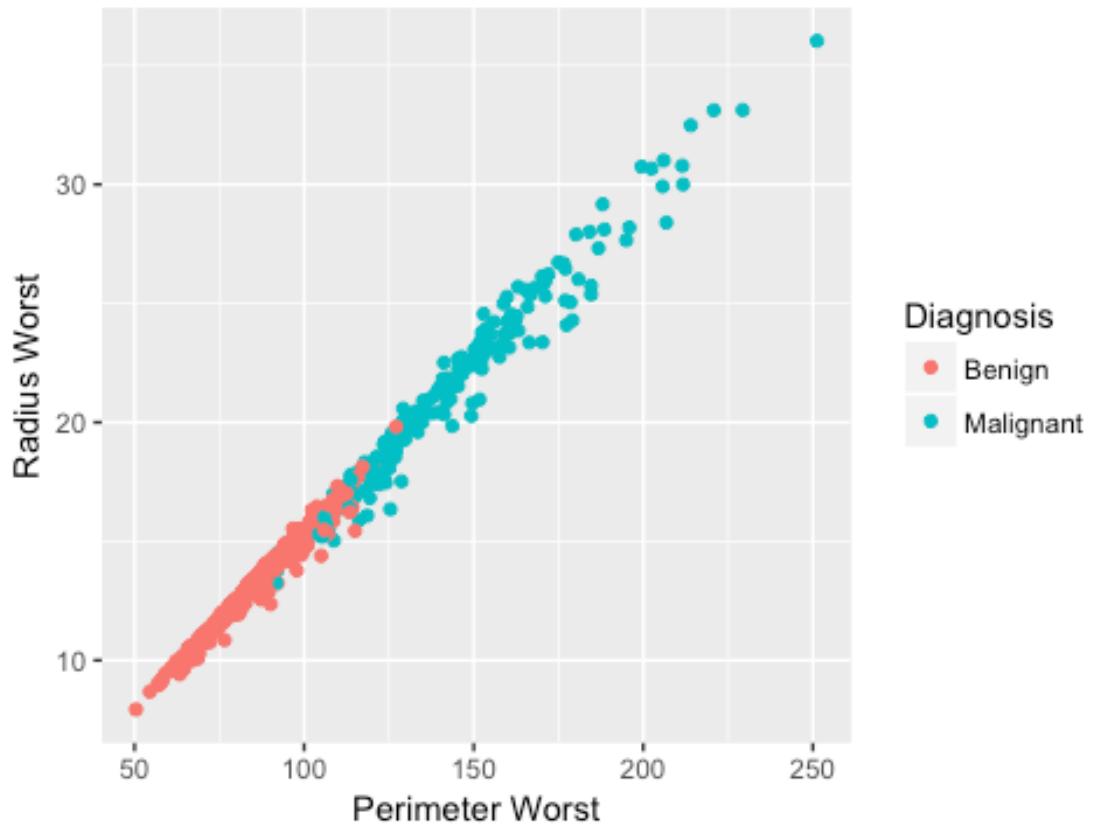
```
#fractal_dimension_worst vs. texture_worst
texture_fractal_dimension_worst_scatter <- ggplot(data = cancer, aes(x = texture_worst, y = fractal_dimension_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Worst", y = "Fractal Dimension Worst", title = "Fractal Dimension Worst vs. Texture Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_fractal_dimension_worst_scatter
```

Fractal Dimension Worst vs. Texture Worst



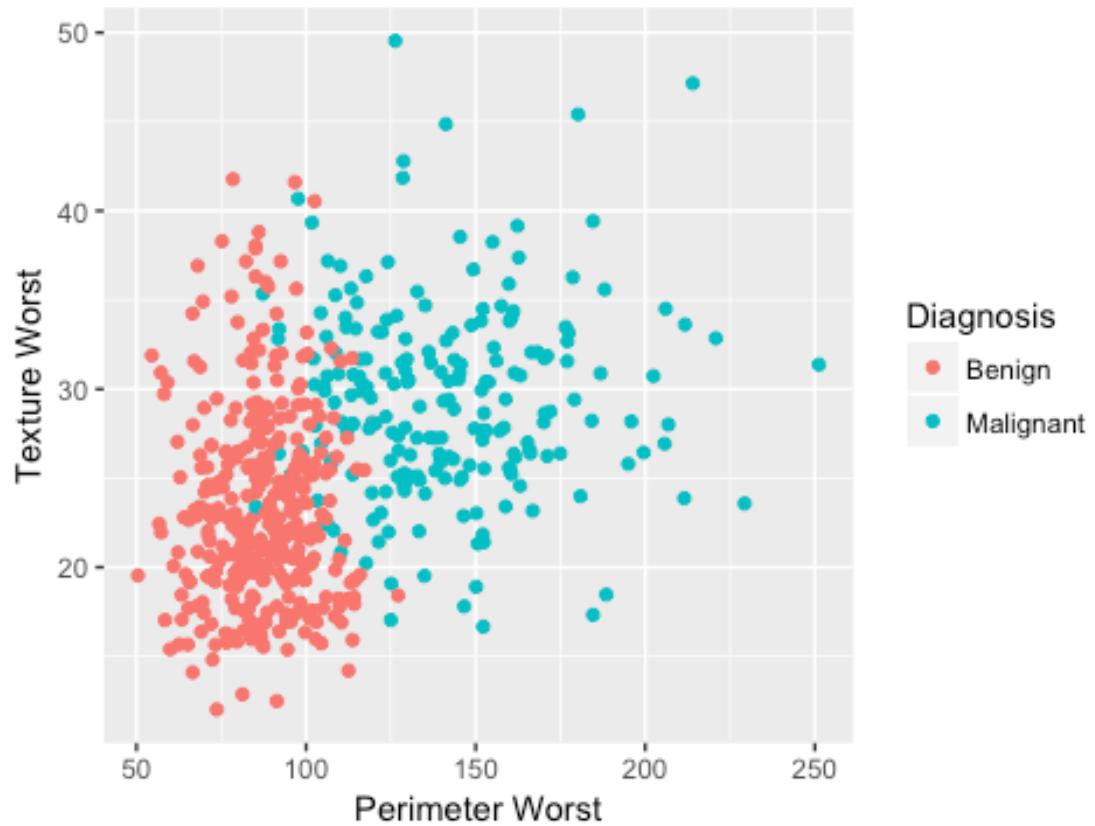
```
#radius_worst vs. perimeter_worst
perimeter_radius_worst_scatter <- ggplot(data = cancer, aes(x = perimeter_worst, y = radius_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Worst", y = "Radius Worst", title = "Radius Worst vs. Perimeter Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_radius_worst_scatter
```

Radius Worst vs. Perimeter Worst



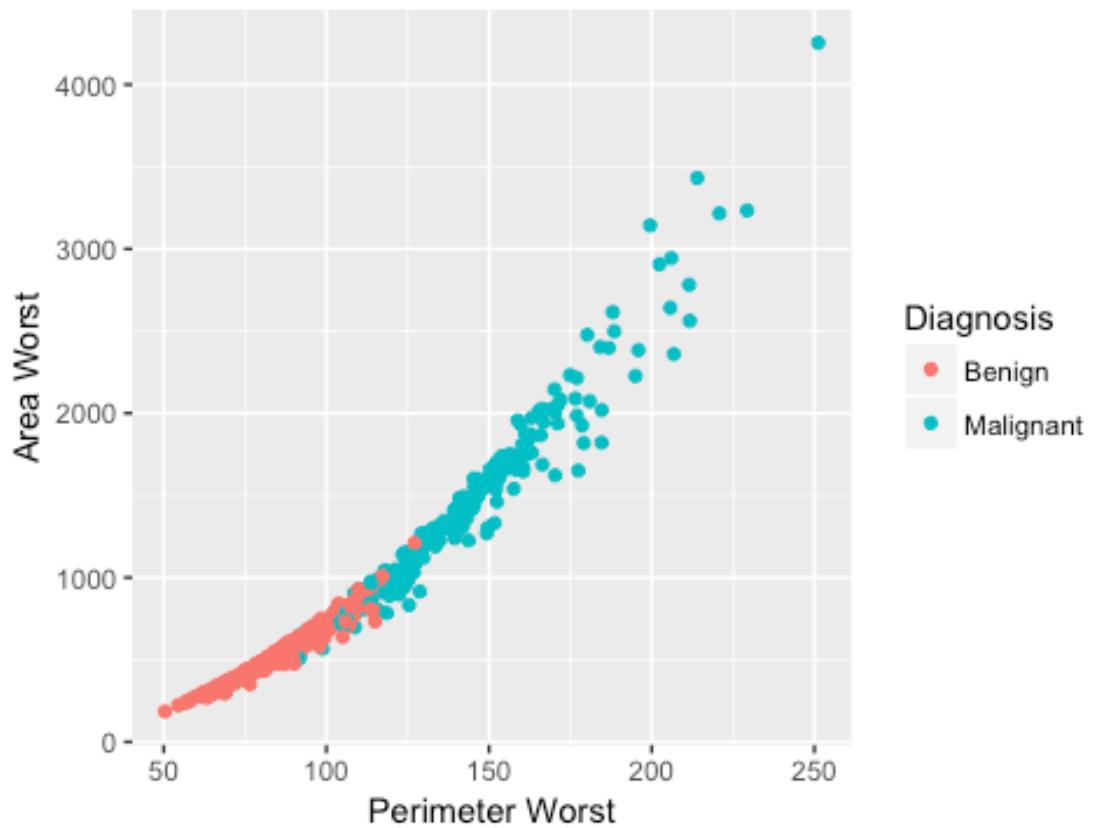
```
#texture_worst vs. perimeter_worst
perimeter_texture_worst_scatter <- ggplot(data = cancer, aes(x = perimeter_worst, y = texture_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Worst", y = "Texture Worst", title = "Texture Worst vs. Perimeter Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_texture_worst_scatter
```

Texture Worst vs. Perimeter Worst



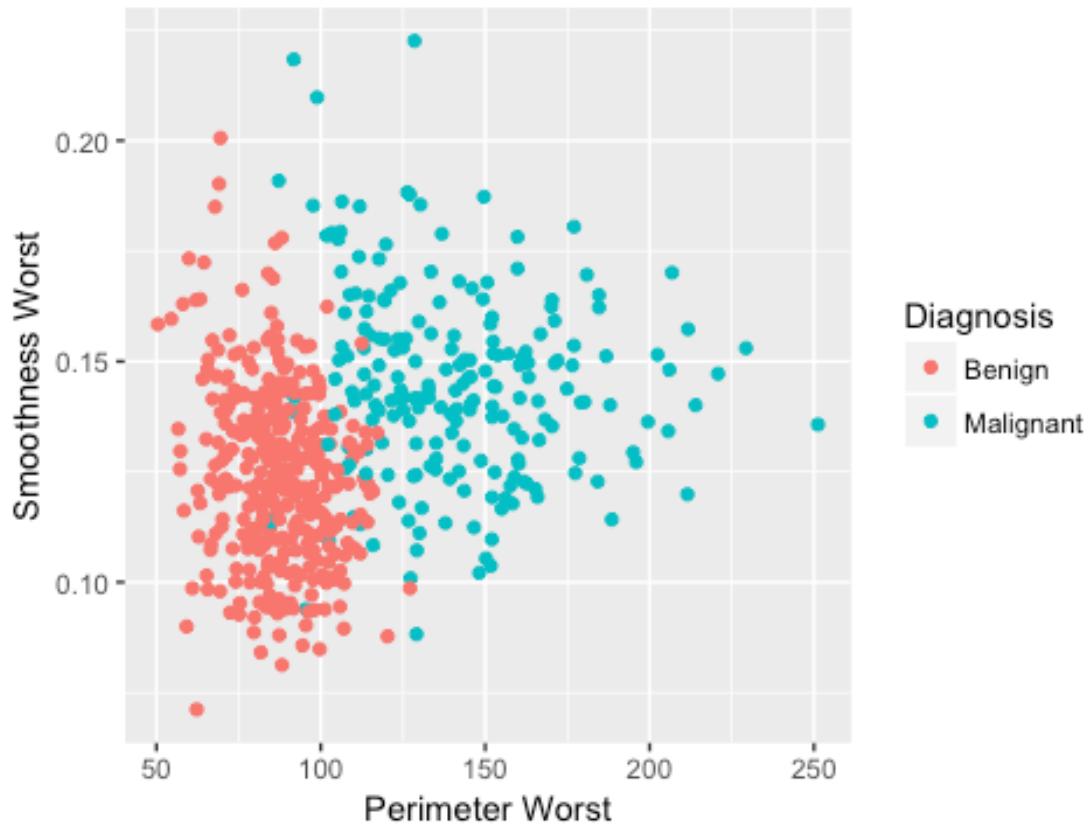
```
#area_worst vs. perimeter_worst
perimeter_area_worst_scatter <- ggplot(data = cancer, aes(x = perimeter_worst,
, y = area_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Worst", y = "Area Worst", title = "Area
Worst vs. Perimeter Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_area_worst_scatter
```

Area Worst vs. Perimeter Worst



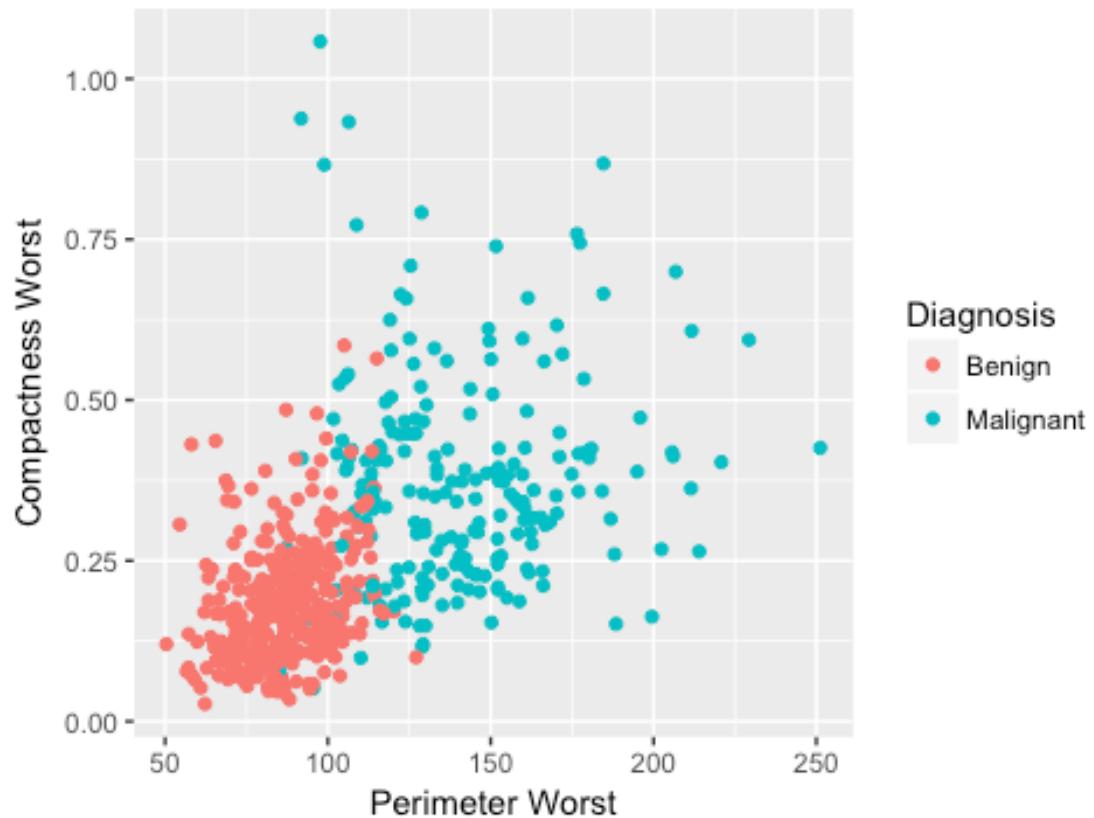
```
#smoothness_worst vs. perimeter_worst
perimeter_smoothness_worst_scatter <- ggplot(data = cancer, aes(x = perimeter_worst, y = smoothness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Worst", y = "Smoothness Worst", title =
"Smoothness Worst vs. Perimeter Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_smoothness_worst_scatter
```

Smoothness Worst vs. Perimeter Worst



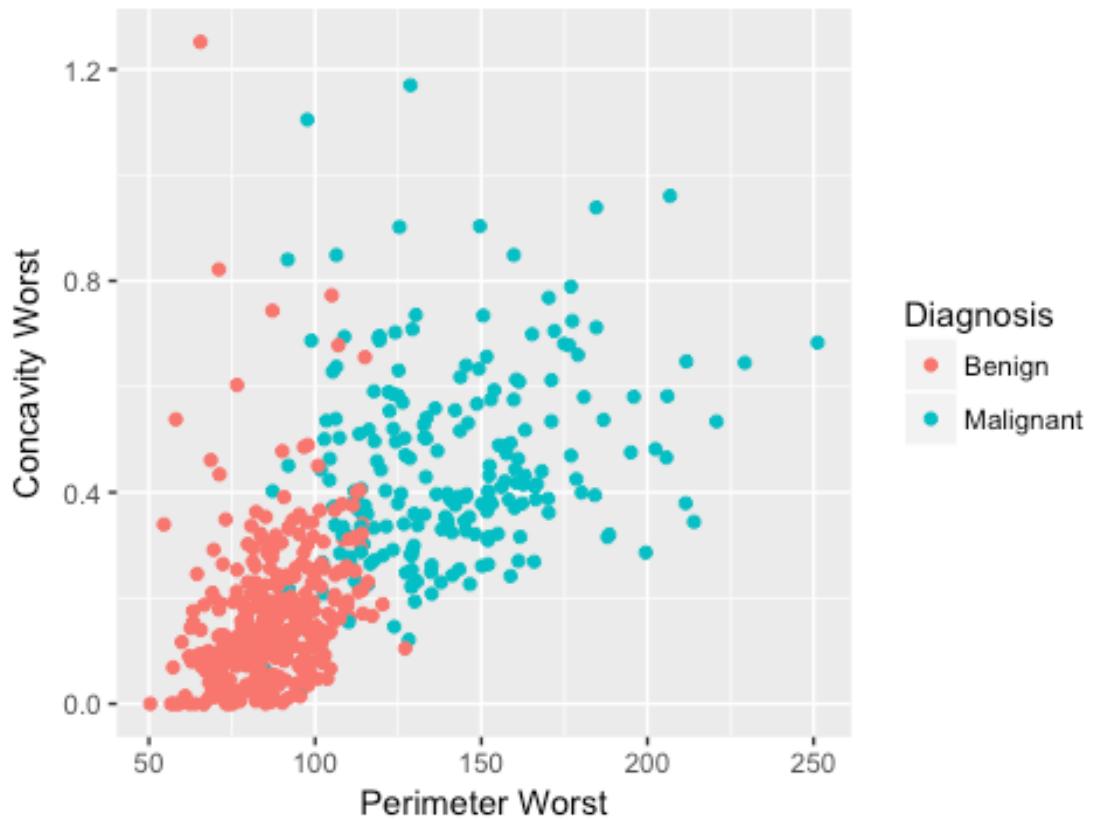
```
#compactness_worst vs. perimeter_worst
perimeter_compactness_worst_scatter <- ggplot(data = cancer, aes(x = perimeter_worst, y = compactness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Worst", y = "Compactness Worst", title =
  "Compactness Worst vs. Perimeter Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_compactness_worst_scatter
```

Compactness Worst vs. Perimeter Worst

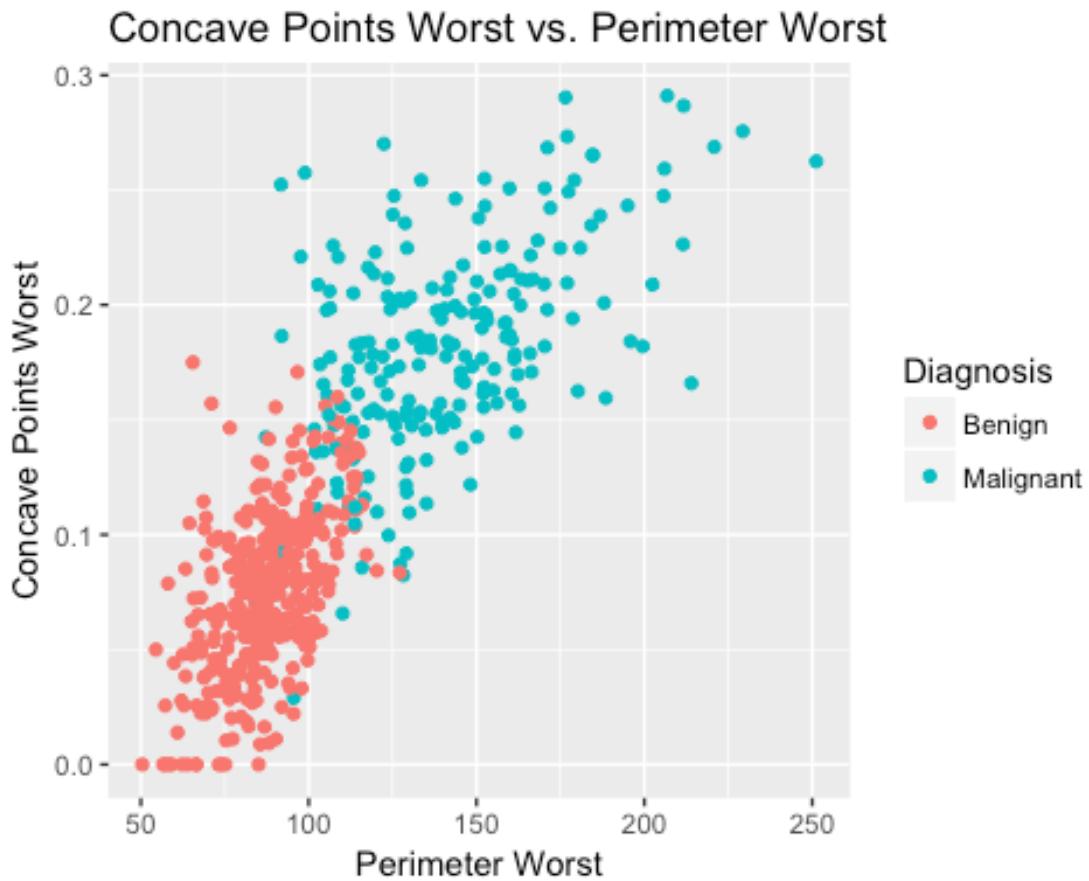


```
#concavity_worst vs. perimeter_worst
perimeter_concavity_worst_scatter <- ggplot(data = cancer, aes(x = perimeter_worst, y = concavity_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Worst", y = "Concavity Worst", title = "Concavity Worst vs. Perimeter Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_concavity_worst_scatter
```

Concavity Worst vs. Perimeter Worst

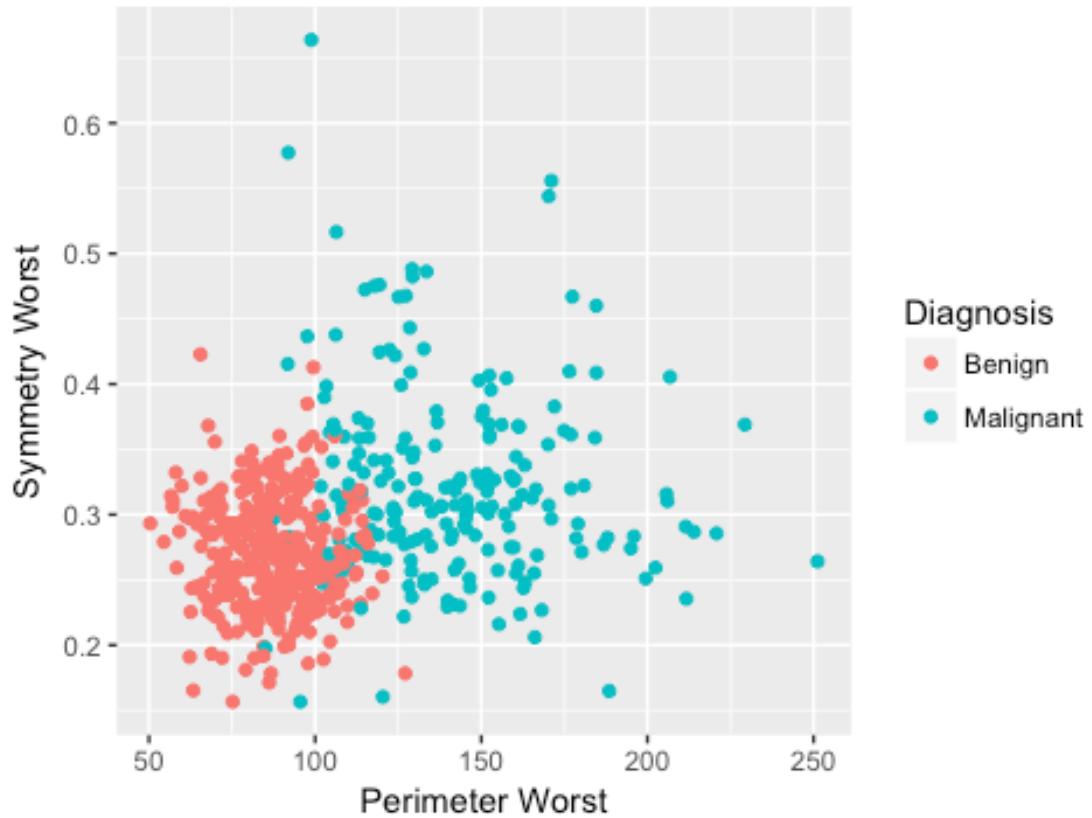


```
#concave.points_worst vs. perimeter_worst
perimeter_concave.points_worst_scatter <- ggplot(data = cancer, aes(x = perimeter_worst, y = concave.points_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Worst", y = "Concave Points Worst", title = "Concave Points Worst vs. Perimeter Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_concave.points_worst_scatter
```



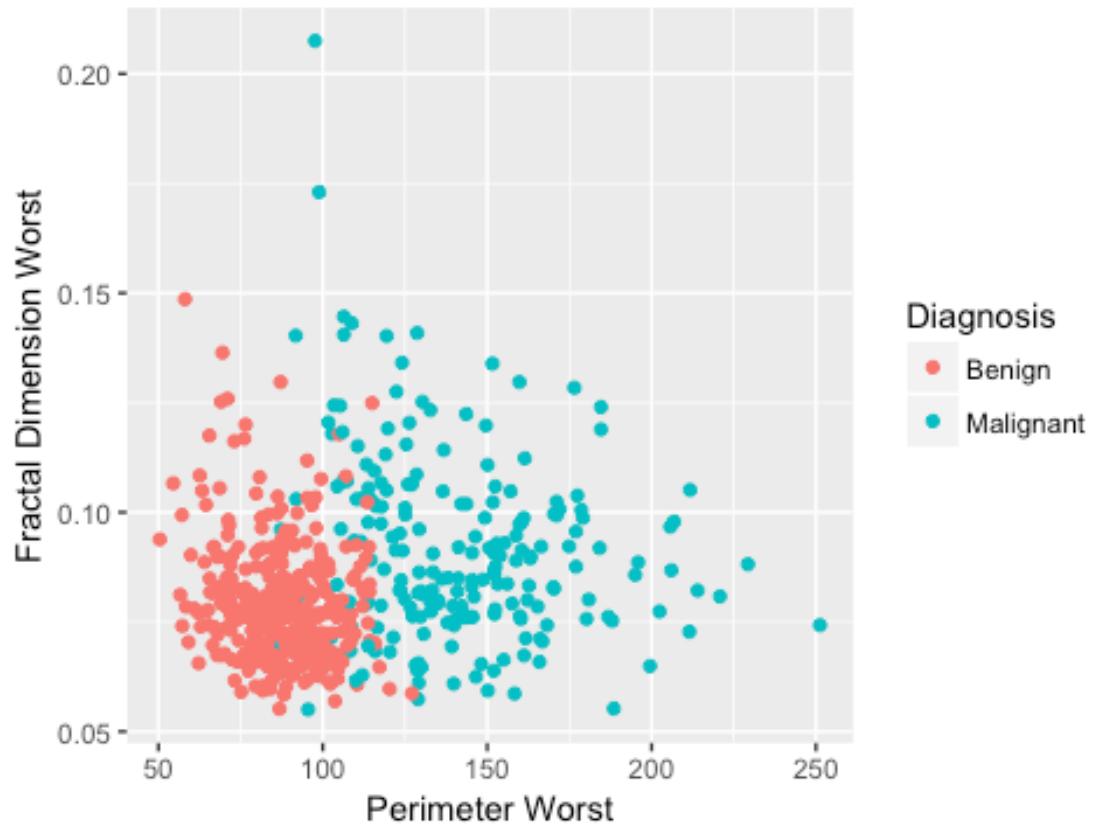
```
#symmetry_worst vs. perimeter_worst
perimeter_symmetry_worst_scatter <- ggplot(data = cancer, aes(x = perimeter_worst, y = symmetry_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Worst", y = "Symmetry Worst", title = "Symmetry Worst vs. Perimeter Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_symmetry_worst_scatter
```

Symmetry Worst vs. Perimeter Worst



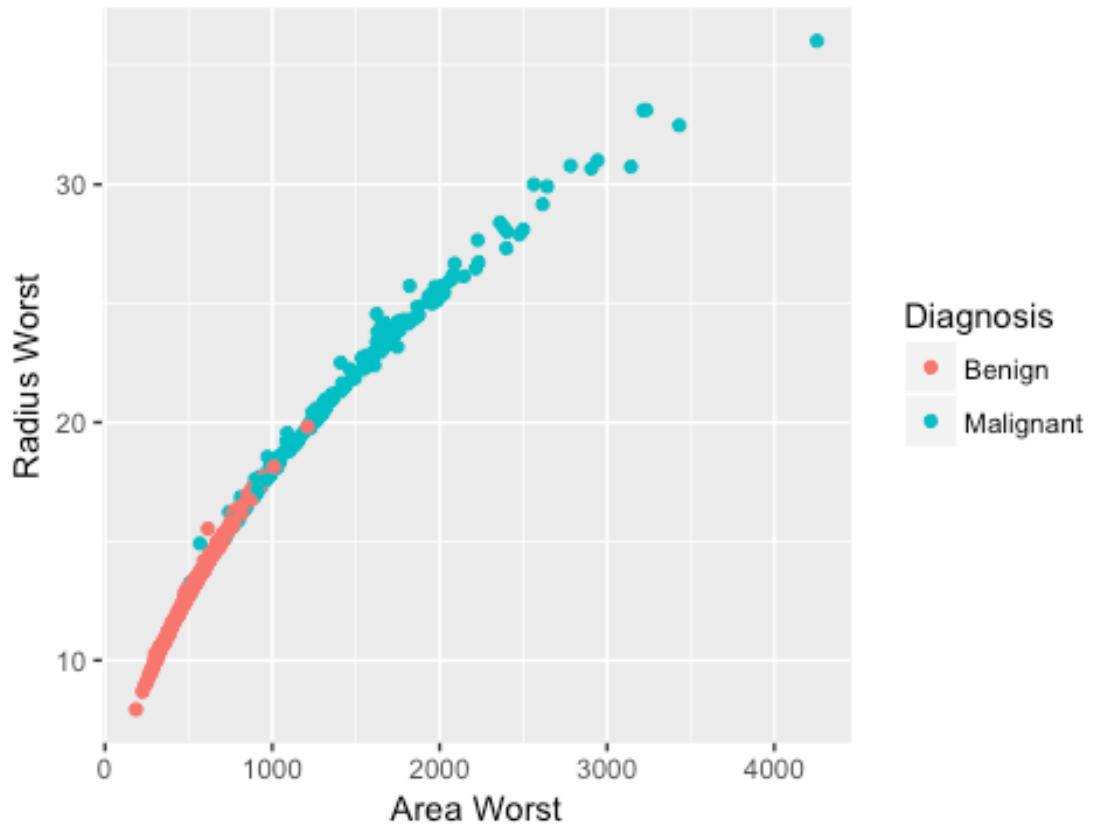
```
#fractal_dimension_worst vs. perimeter_worst
perimeter_fractal_dimension_worst_scatter <- ggplot(data = cancer, aes(x = perimeter_worst, y = fractal_dimension_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Worst", y = "Fractal Dimension Worst", title = "Fractal Dimension Worst vs. Perimeter Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_fractal_dimension_worst_scatter
```

Fractal Dimension Worst vs. Perimeter Worst



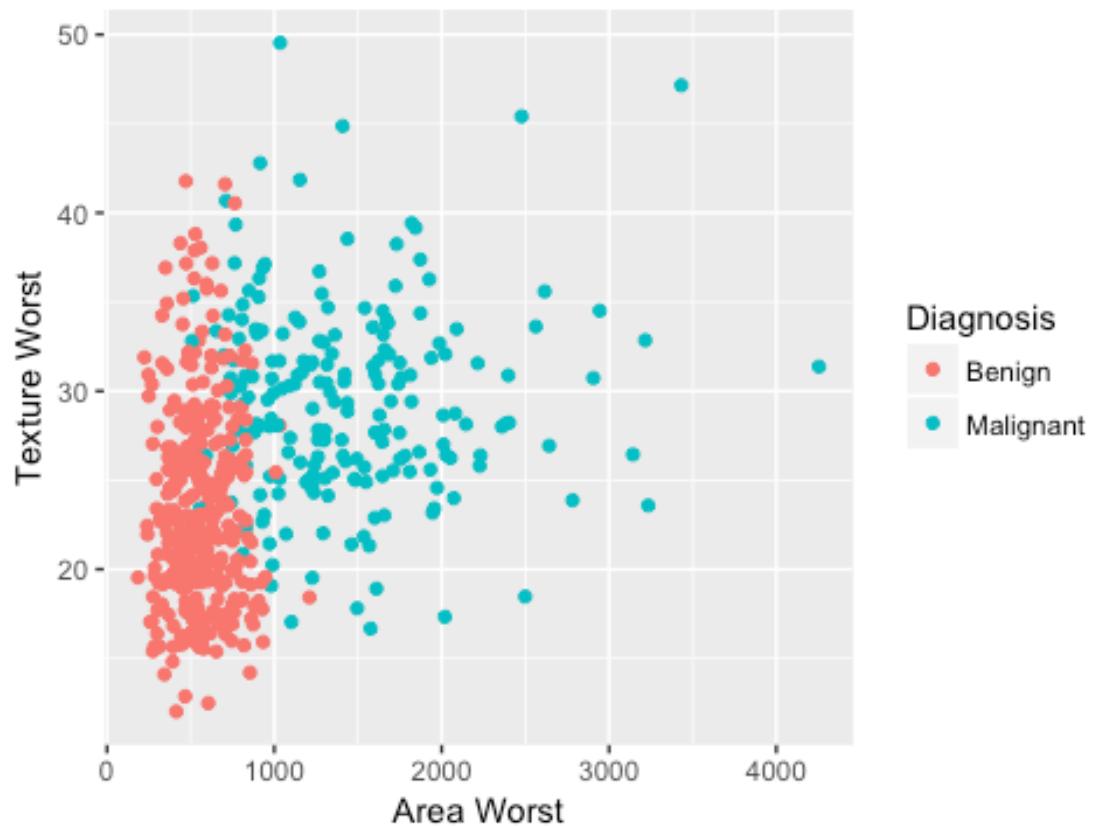
```
#radius_worst vs. area_worst
area_radius_worst_scatter <- ggplot(data = cancer, aes(x = area_worst, y = radius_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Worst", y = "Radius Worst", title = "Radius Worst vs. Area Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_radius_worst_scatter
```

Radius Worst vs. Area Worst



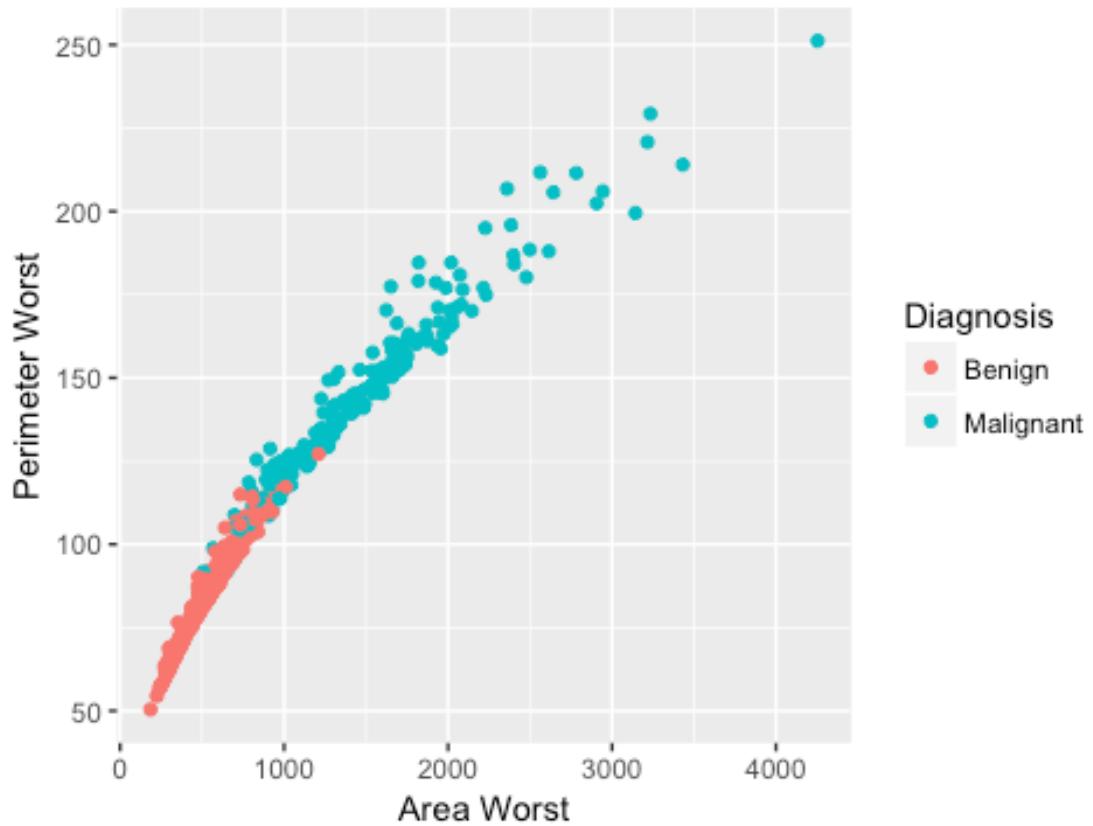
```
#texture_worst vs. area_worst
area_texture_worst_scatter <- ggplot(data = cancer, aes(x = area_worst, y = texture_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Worst", y = "Texture Worst", title = "Texture Worst vs. Area Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_texture_worst_scatter
```

Texture Worst vs. Area Worst



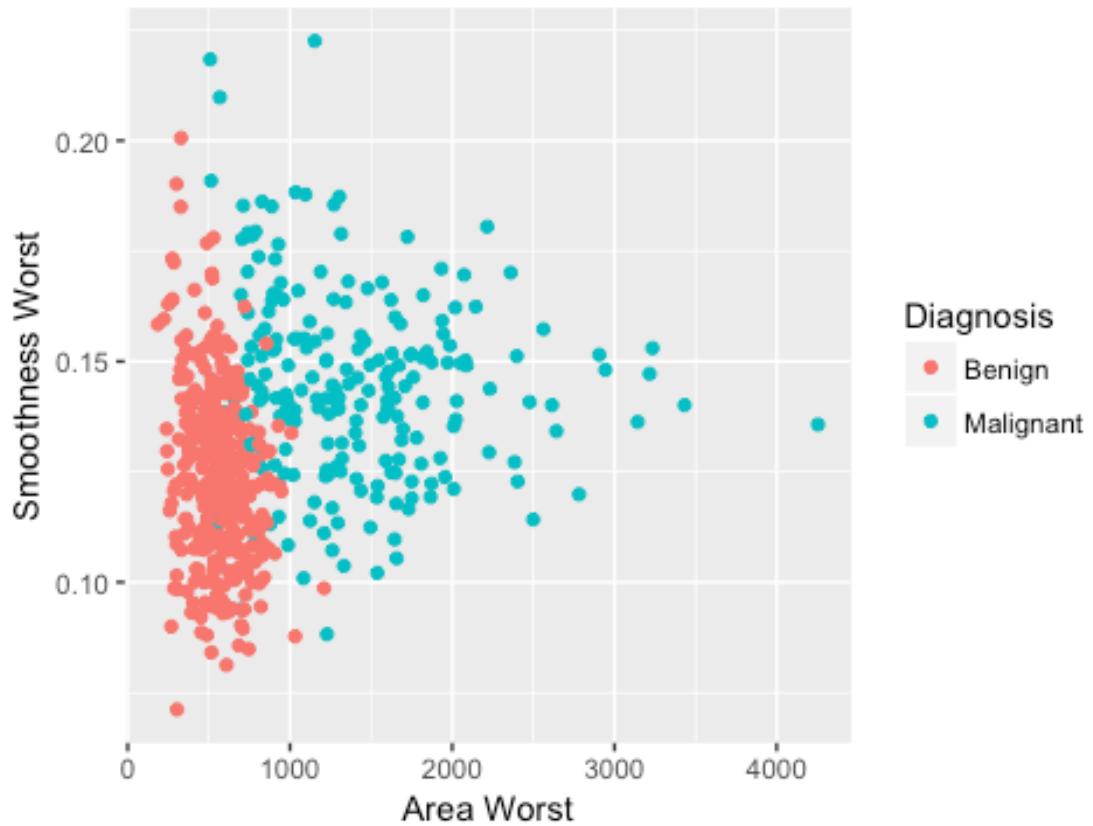
```
#perimeter_worst vs. area_worst
area_perimeter_worst_scatter <- ggplot(data = cancer, aes(x = area_worst, y = perimeter_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Worst", y = "Perimeter Worst", title = "Perimeter Worst vs. Area Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_perimeter_worst_scatter
```

Perimeter Worst vs. Area Worst



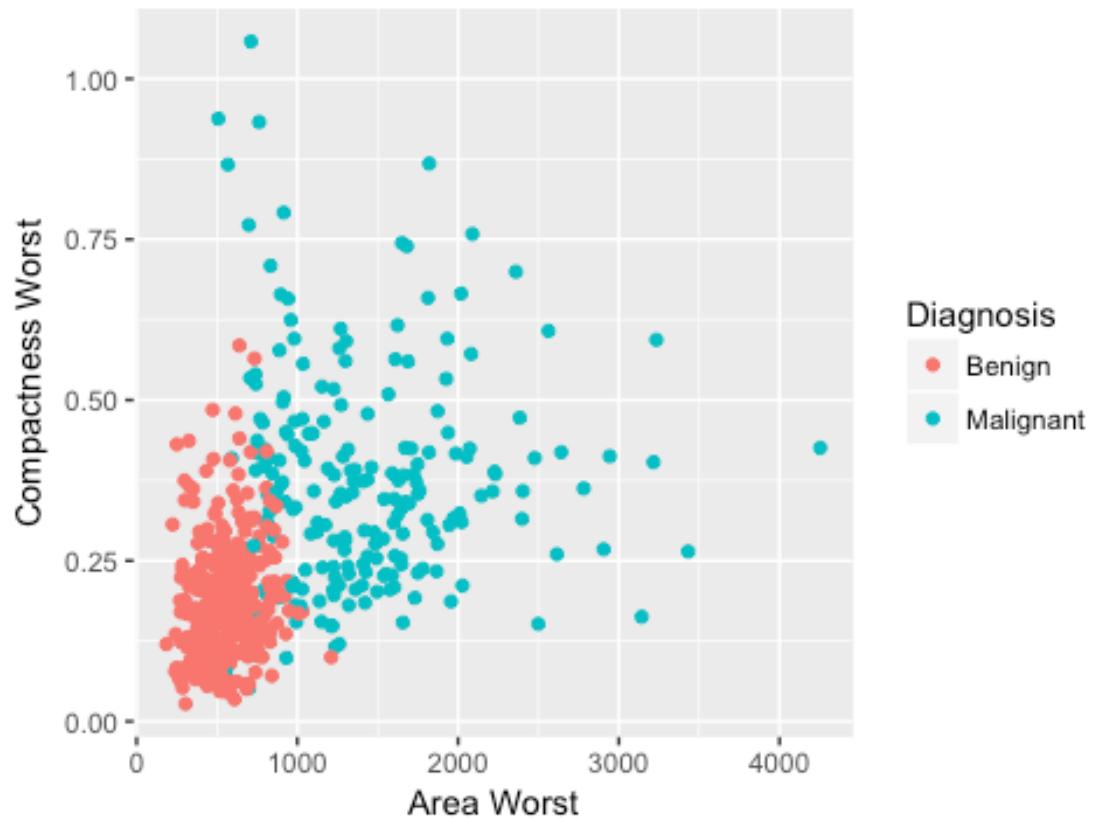
```
#smoothness_worst vs. area_worst
area_smoothness_worst_scatter <- ggplot(data = cancer, aes(x = area_worst, y = smoothness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Worst", y = "Smoothness Worst", title = "Smoothness Worst vs. Area Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_smoothness_worst_scatter
```

Smoothness Worst vs. Area Worst



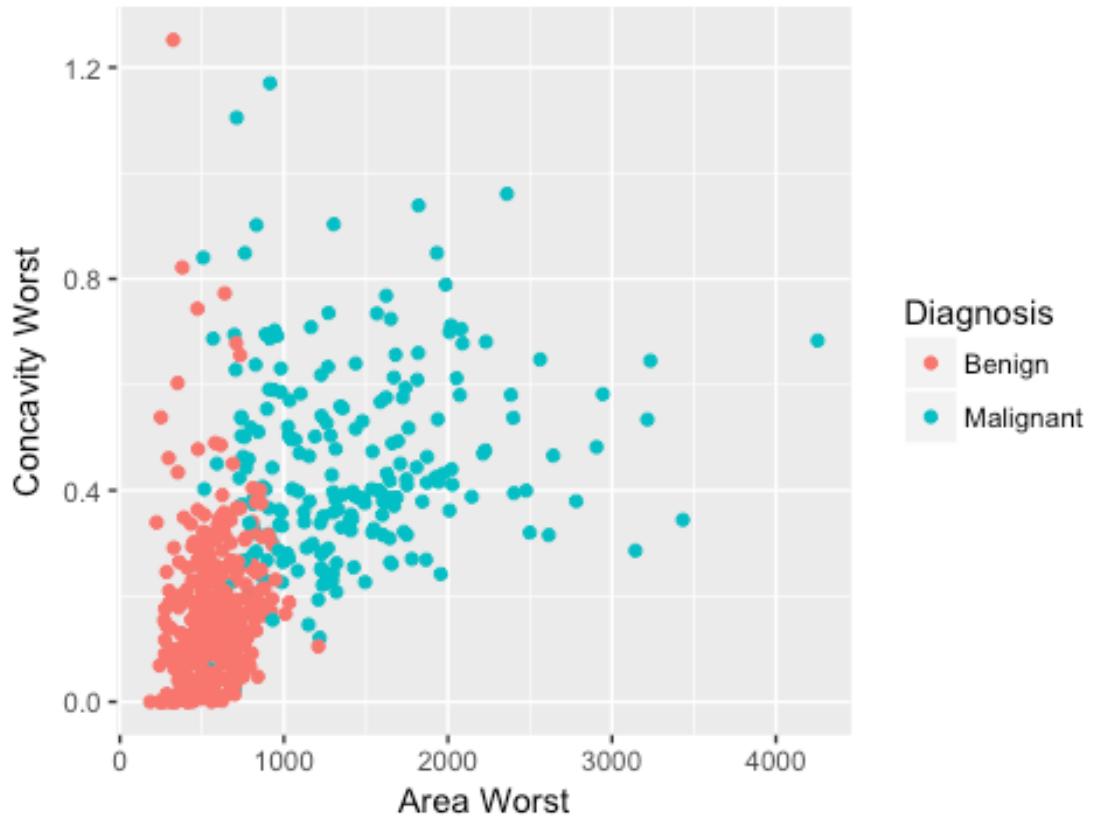
```
#compactness_worst vs. area_worst
area_compactness_worst_scatter <- ggplot(data = cancer, aes(x = area_worst, y = compactness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Worst", y = "Compactness Worst", title = "Compactness Worst vs. Area Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_compactness_worst_scatter
```

Compactness Worst vs. Area Worst

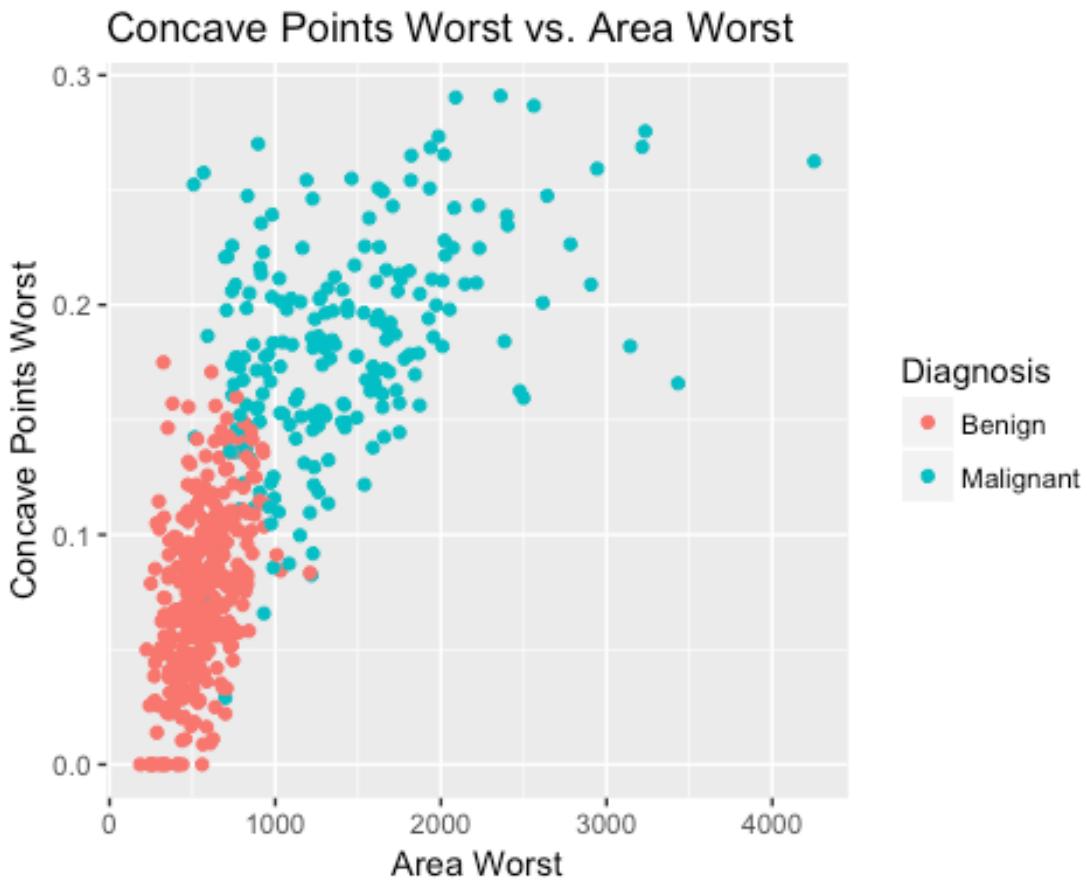


```
#concavity_worst vs. area_worst
area_concavity_worst_scatter <- ggplot(data = cancer, aes(x = area_worst, y = concavity_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Worst", y = "Concavity Worst", title = "Concavity Worst vs. Area Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_concavity_worst_scatter
```

Concavity Worst vs. Area Worst

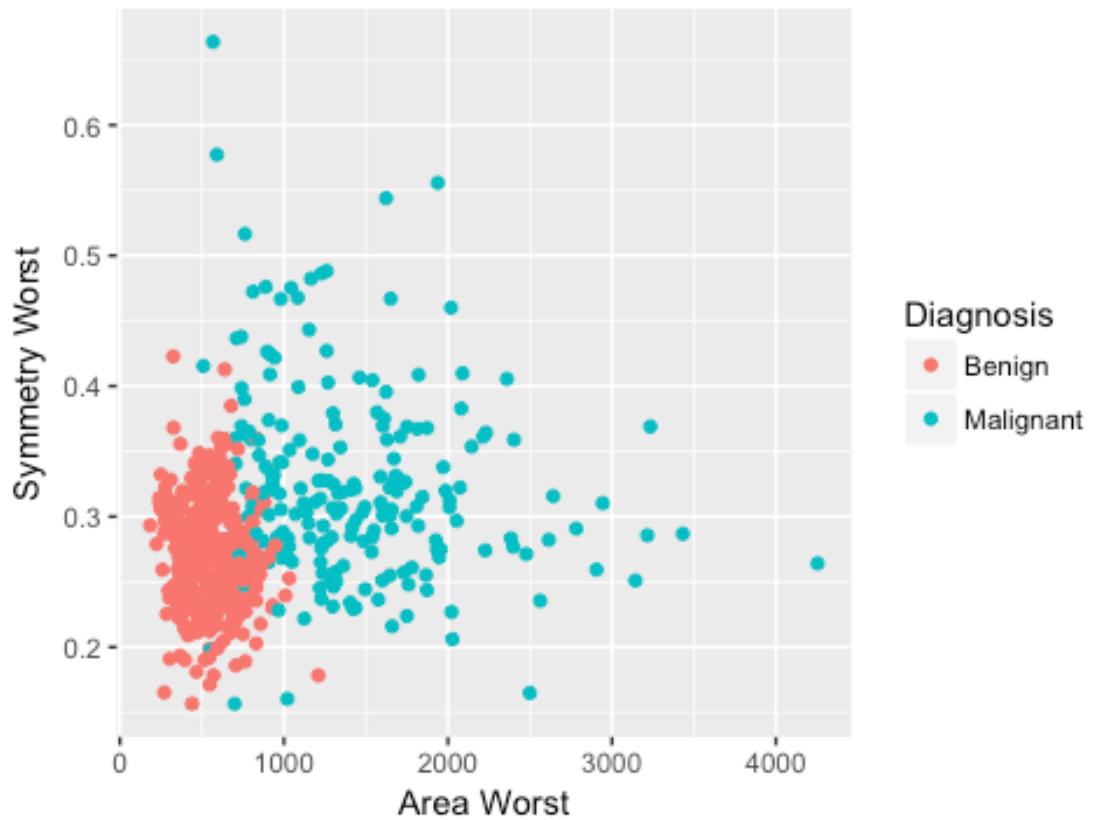


```
#concave.points_worst vs. area_worst
area_concave.points_worst_scatter <- ggplot(data = cancer, aes(x = area_worst,
, y = concave.points_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Worst", y = "Concave Points Worst", title = "Concave Points Worst vs. Area Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_concave.points_worst_scatter
```



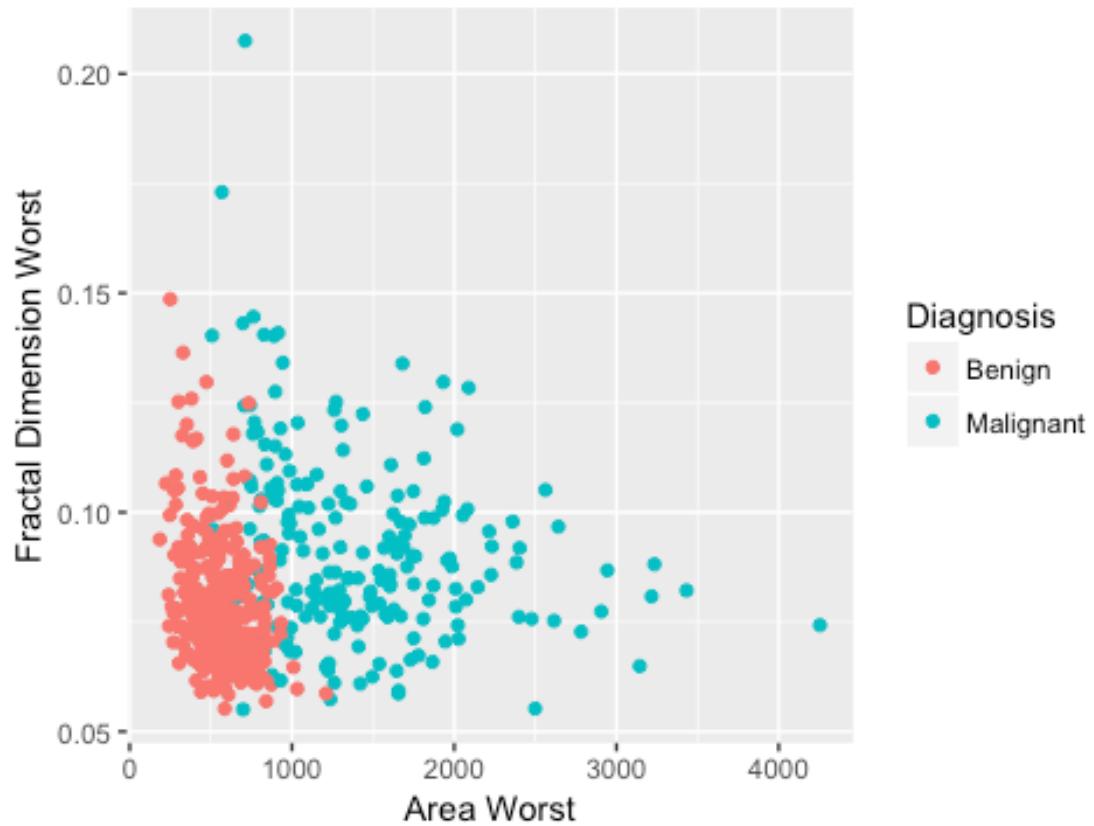
```
#symmetry_worst vs. area_worst
area_symmetry_worst_scatter <- ggplot(data = cancer, aes(x = area_worst, y = symmetry_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Worst", y = "Symmetry Worst", title = "Symmetry Worst vs. Area Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_symmetry_worst_scatter
```

Symmetry Worst vs. Area Worst



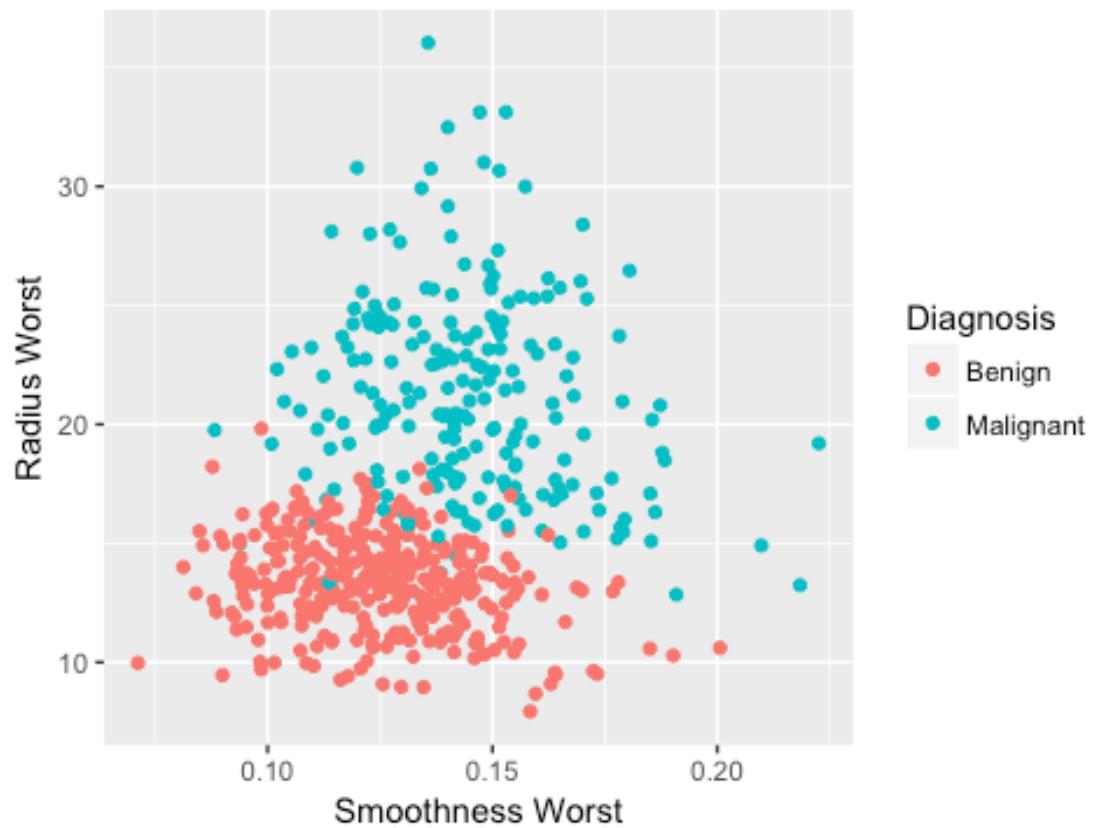
```
#fractal_dimension_worst vs. area_worst
area_fractal_dimension_worst_scatter <- ggplot(data = cancer, aes(x = area_worst, y = fractal_dimension_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Worst", y = "Fractal Dimension Worst", title =
  "Fractal Dimension Worst vs. Area Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_fractal_dimension_worst_scatter
```

Fractal Dimension Worst vs. Area Worst



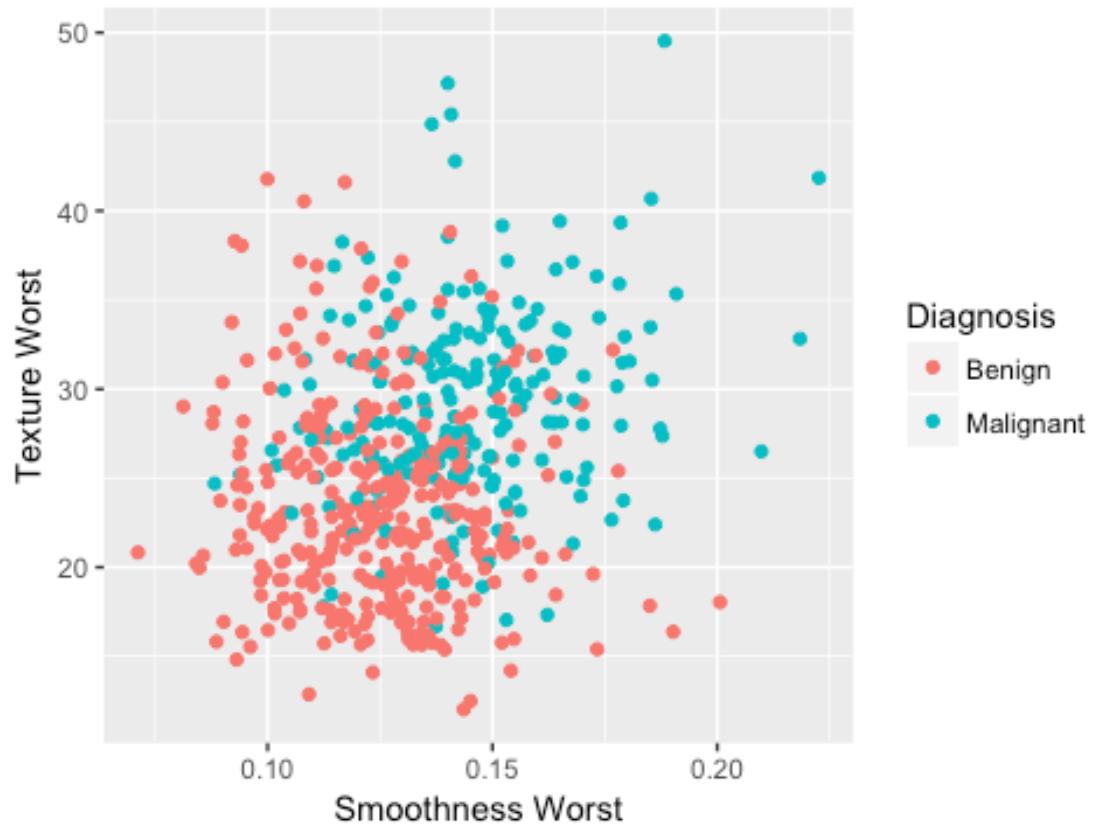
```
#radius_worst vs. smoothness_worst
smoothness_radius_worst_scatter <- ggplot(data = cancer, aes(x = smoothness_worst, y = radius_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Worst", y = "Radius Worst", title = "Radius Worst vs. Smoothness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_radius_worst_scatter
```

Radius Worst vs. Smoothness Worst



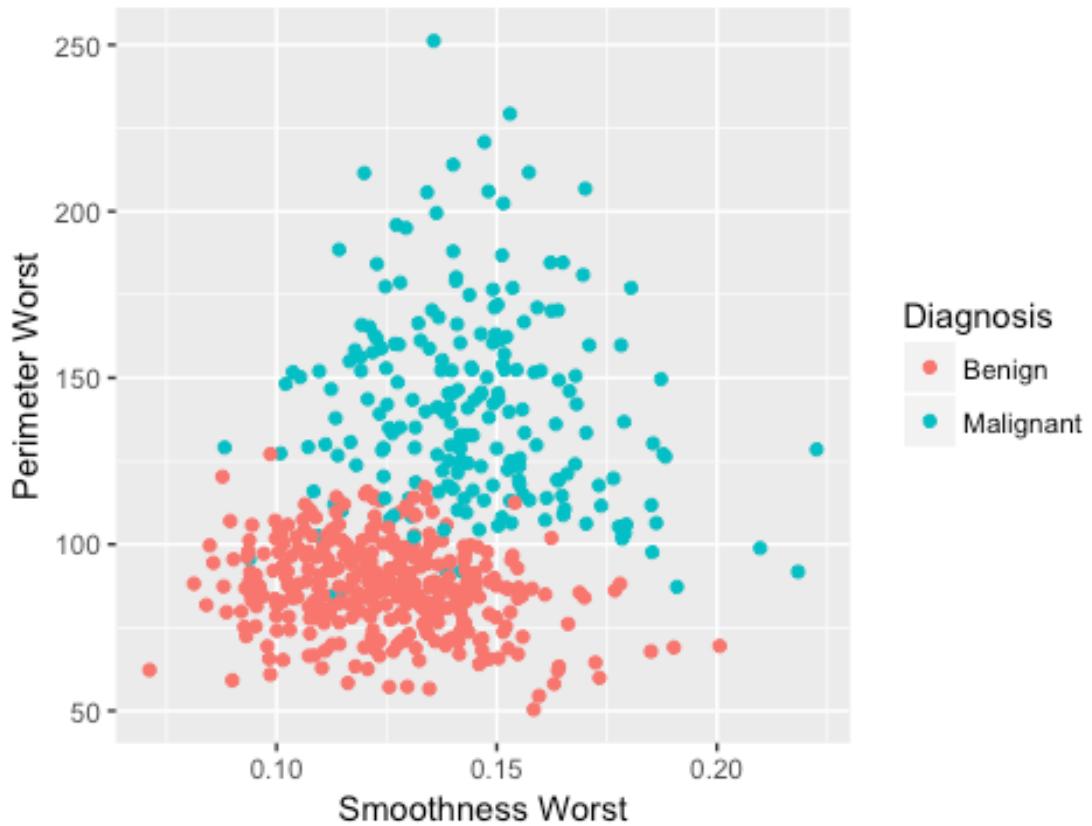
```
#texture_worst vs. smoothness_worst
smoothness_texture_worst_scatter <- ggplot(data = cancer, aes(x = smoothness_worst, y = texture_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Worst", y = "Texture Worst", title = "Texture Worst vs. Smoothness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_texture_worst_scatter
```

Texture Worst vs. Smoothness Worst



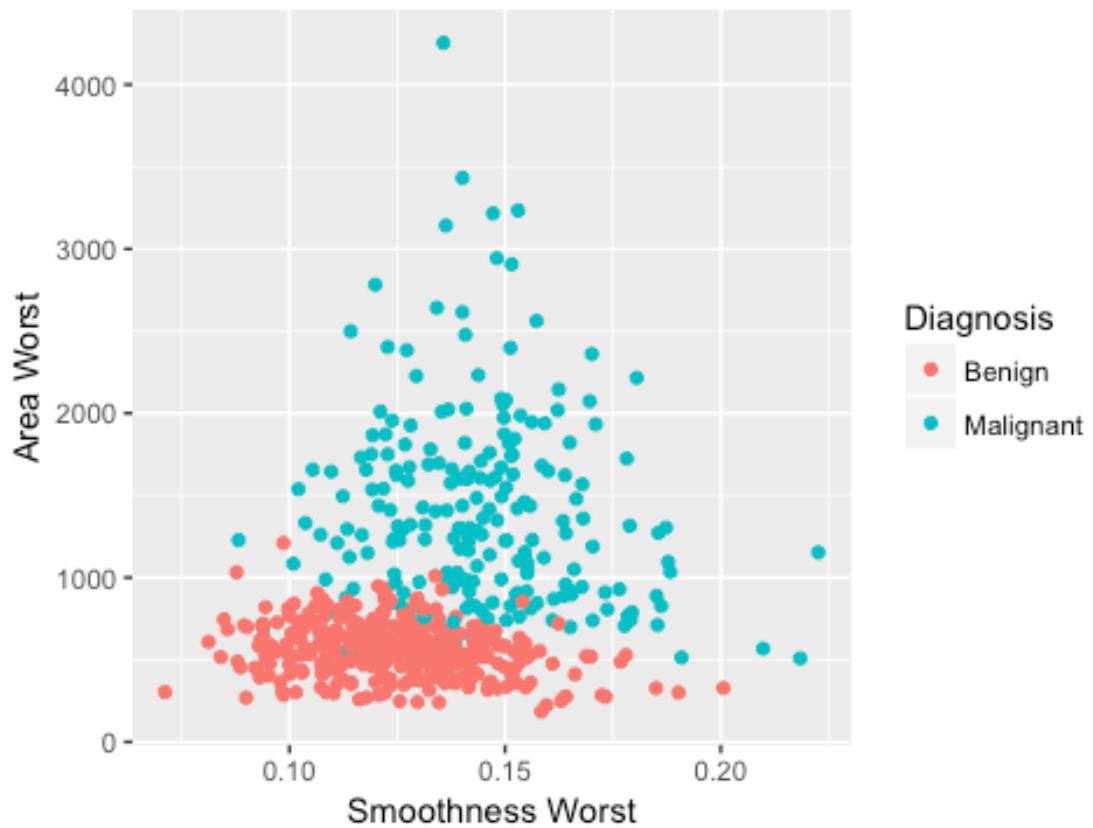
```
#perimeter_worst vs. smoothness_worst
smoothness_perimeter_worst_scatter <- ggplot(data = cancer, aes(x = smoothness_worst, y = perimeter_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Worst", y = "Perimeter Worst", title =
  "Perimeter Worst vs. Smoothness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_perimeter_worst_scatter
```

Perimeter Worst vs. Smoothness Worst



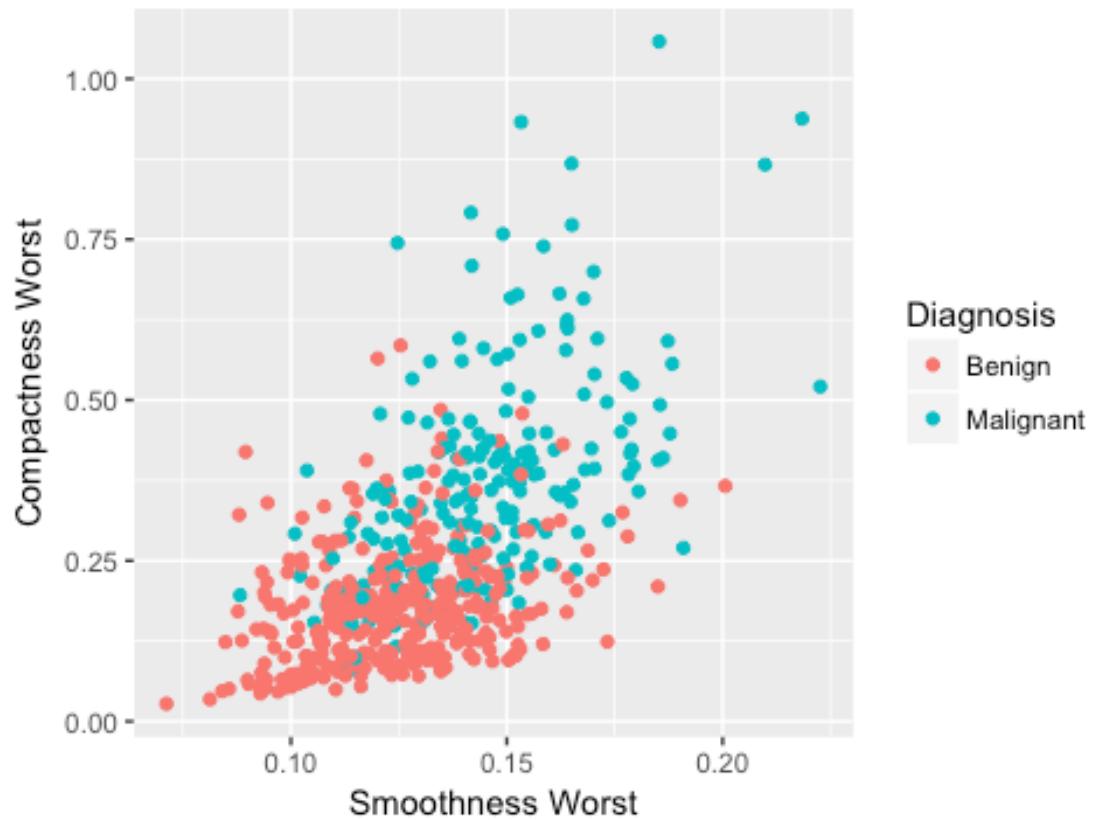
```
#area_worst vs. smoothness_worst
smoothness_area_worst_scatter <- ggplot(data = cancer, aes(x = smoothness_worst, y = area_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Worst", y = "Area Worst", title = "Area Worst vs. Smoothness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_area_worst_scatter
```

Area Worst vs. Smoothness Worst



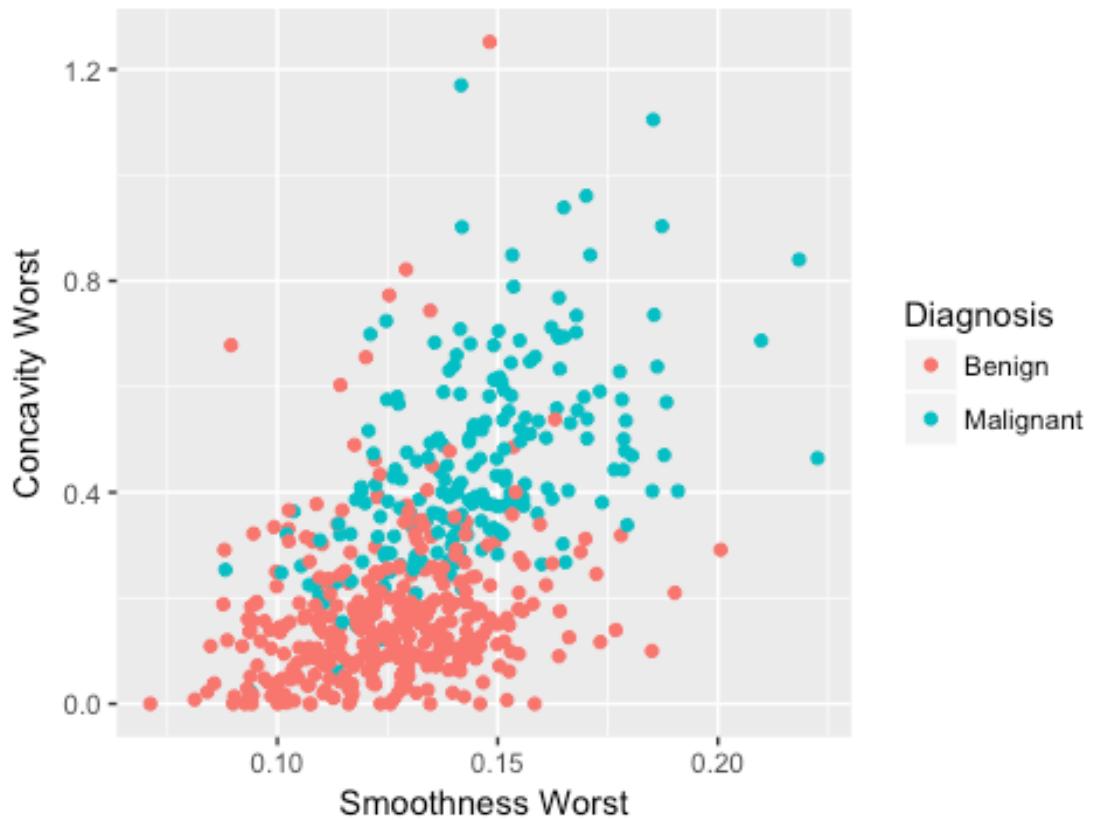
```
#compactness_worst vs. smoothness_worst
smoothness_compactness_worst_scatter <- ggplot(data = cancer, aes(x = smoothness_worst, y = compactness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Worst", y = "Compactness Worst", title =
  "Compactness Worst vs. Smoothness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_compactness_worst_scatter
```

Compactness Worst vs. Smoothness Worst



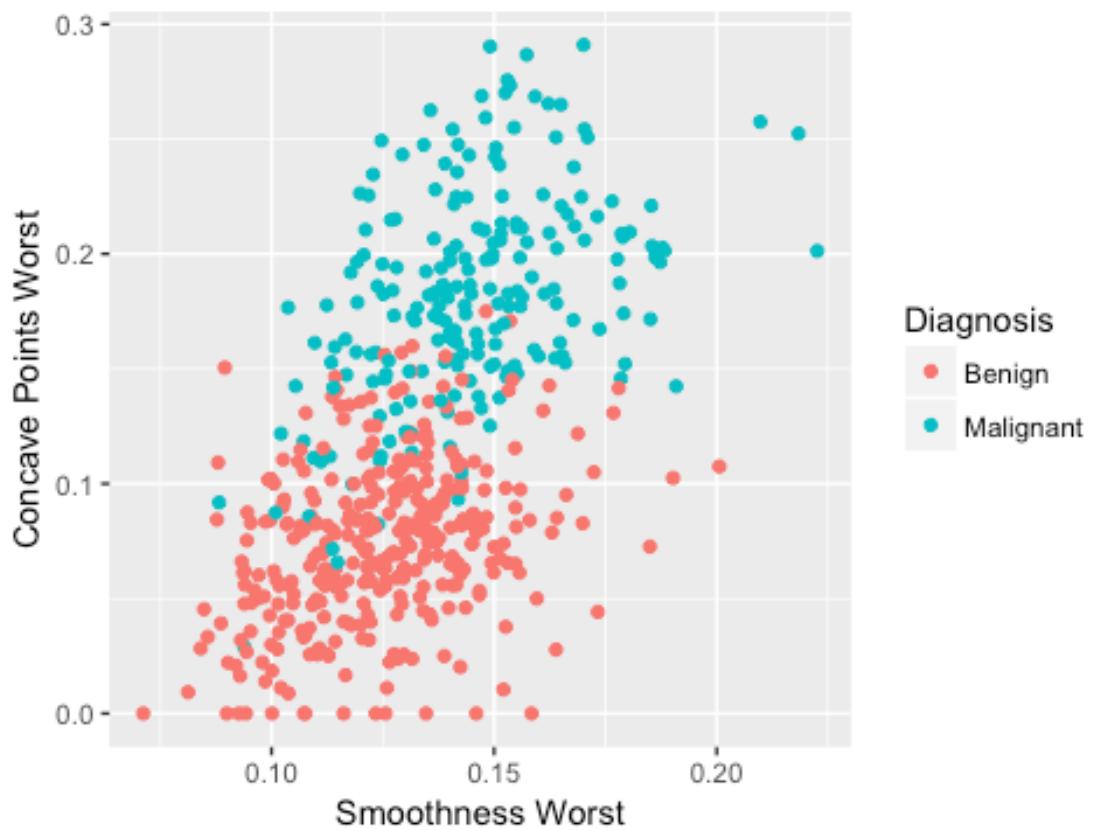
```
#concavity_worst vs. smoothness_worst
smoothness_concavity_worst_scatter <- ggplot(data = cancer, aes(x = smoothness_worst, y = concavity_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Worst", y = "Concavity Worst", title =
  "Concavity Worst vs. Smoothness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_concavity_worst_scatter
```

Concavity Worst vs. Smoothness Worst



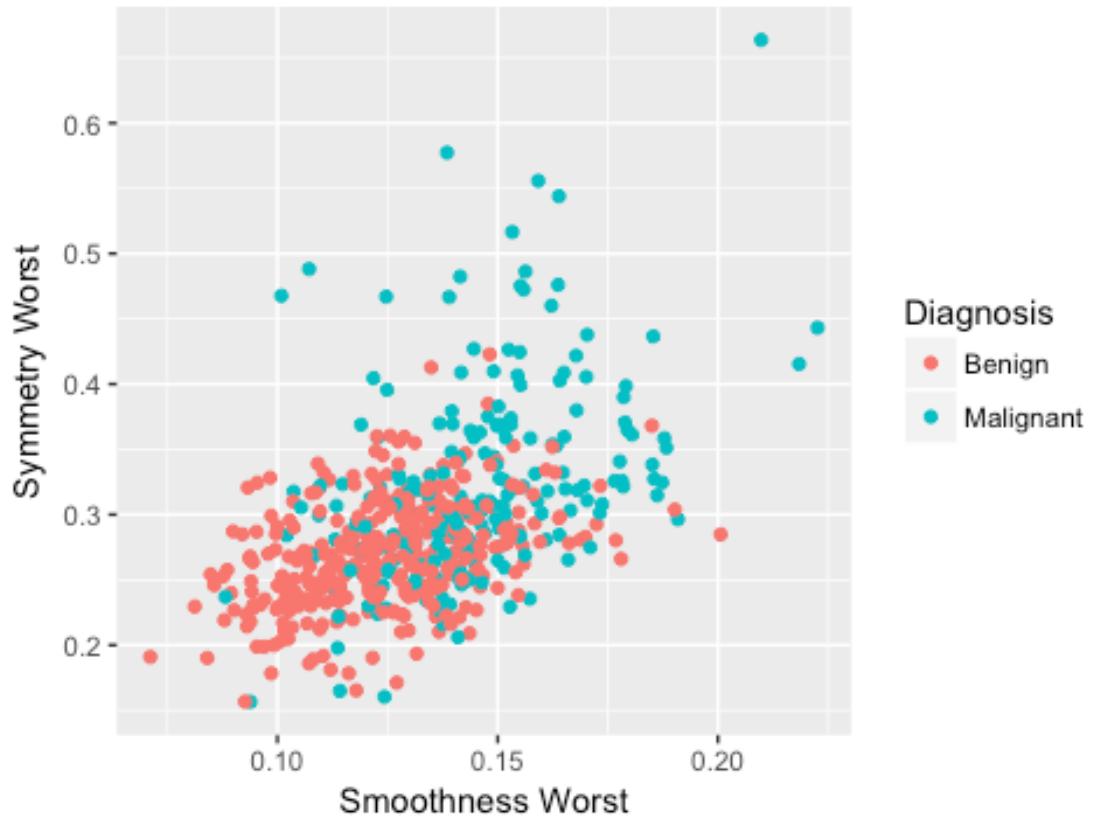
```
#concave.points_worst vs. smoothness_worst
smoothness_concave.points_worst_scatter <- ggplot(data = cancer, aes(x = smoothness_worst, y = concave.points_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Worst", y = "Concave Points Worst", title = "Concave Points Worst vs. Smoothness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_concave.points_worst_scatter
```

Concave Points Worst vs. Smoothness Worst



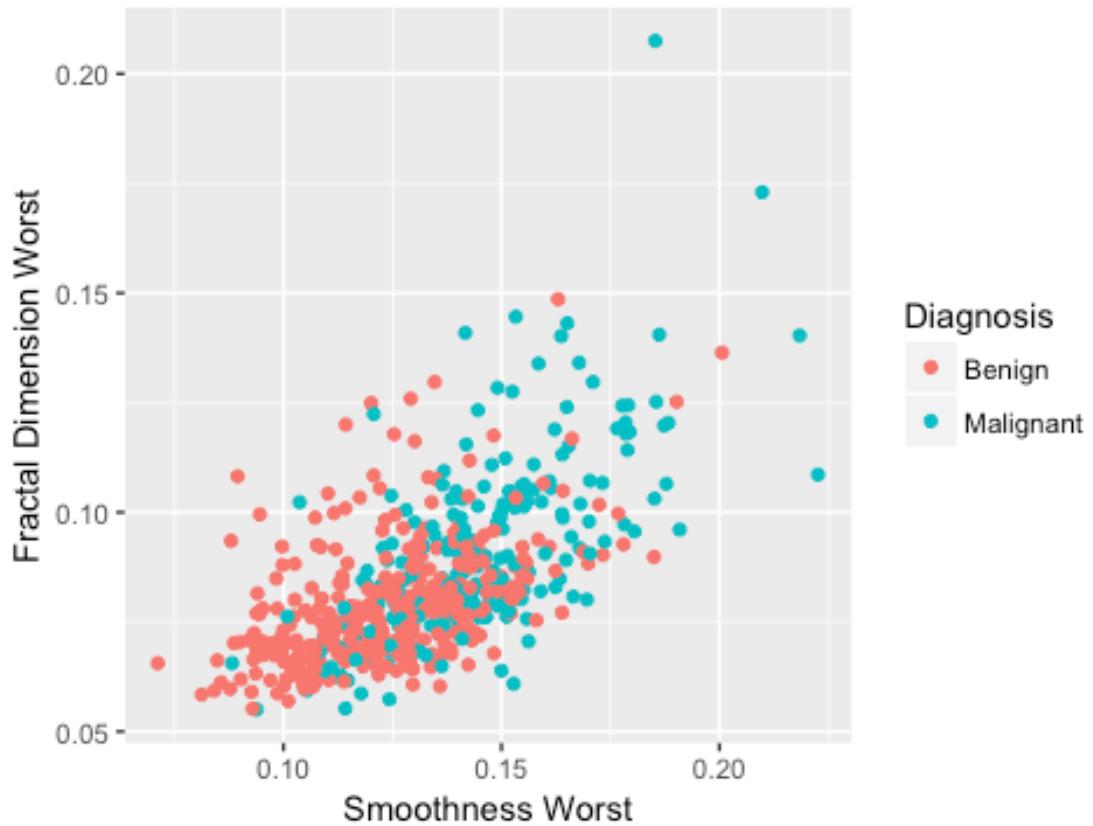
```
#symmetry_worst vs. smoothness_worst
smoothness_symmetry_worst_scatter <- ggplot(data = cancer, aes(x = smoothness_worst, y = symmetry_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Worst", y = "Symmetry Worst", title = "Symmetry Worst vs. Smoothness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_symmetry_worst_scatter
```

Symmetry Worst vs. Smoothness Worst



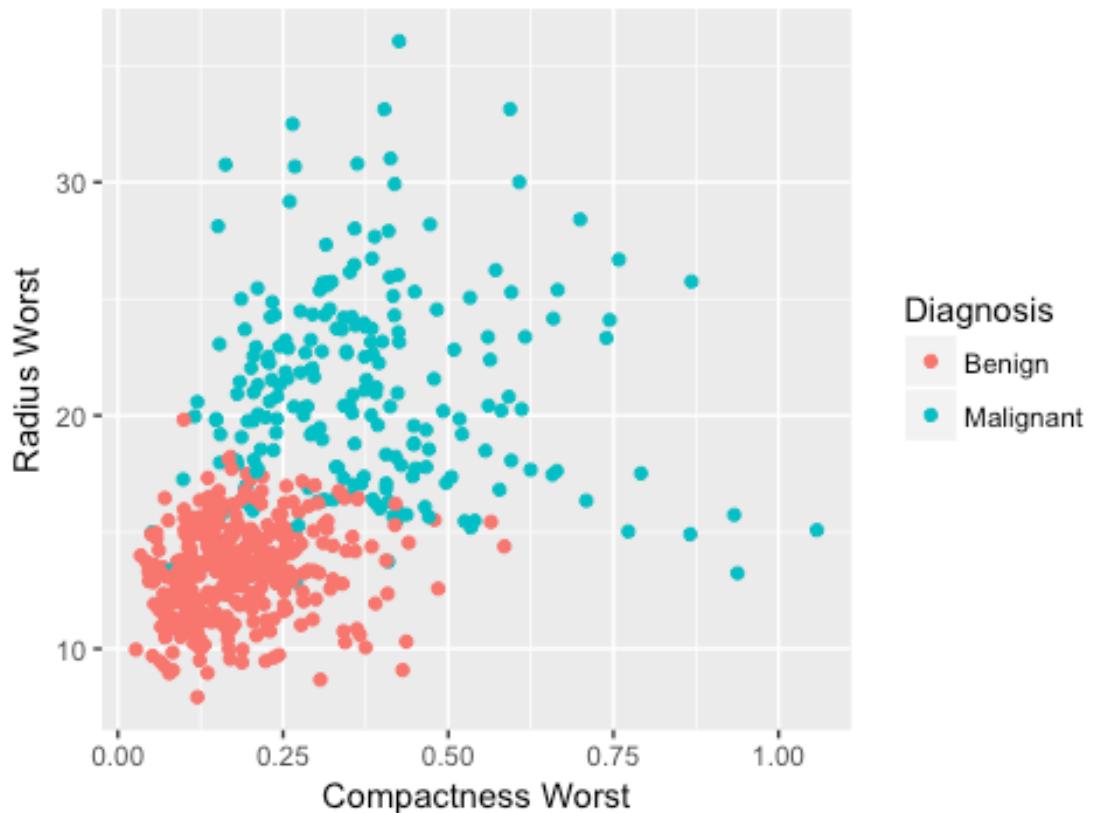
```
#fractal_dimension_worst vs. smoothness_worst
smoothness_fractal_dimension_worst_scatter <- ggplot(data = cancer, aes(x = smoothness_worst, y = fractal_dimension_worst, color = as.factor(diagnosis)))
+
  geom_point() + labs(x = "Smoothness Worst", y = "Fractal Dimension Worst",
  title = "Fractal Dimension Worst vs. Smoothness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_fractal_dimension_worst_scatter
```

Fractal Dimension Worst vs. Smoothness Worst



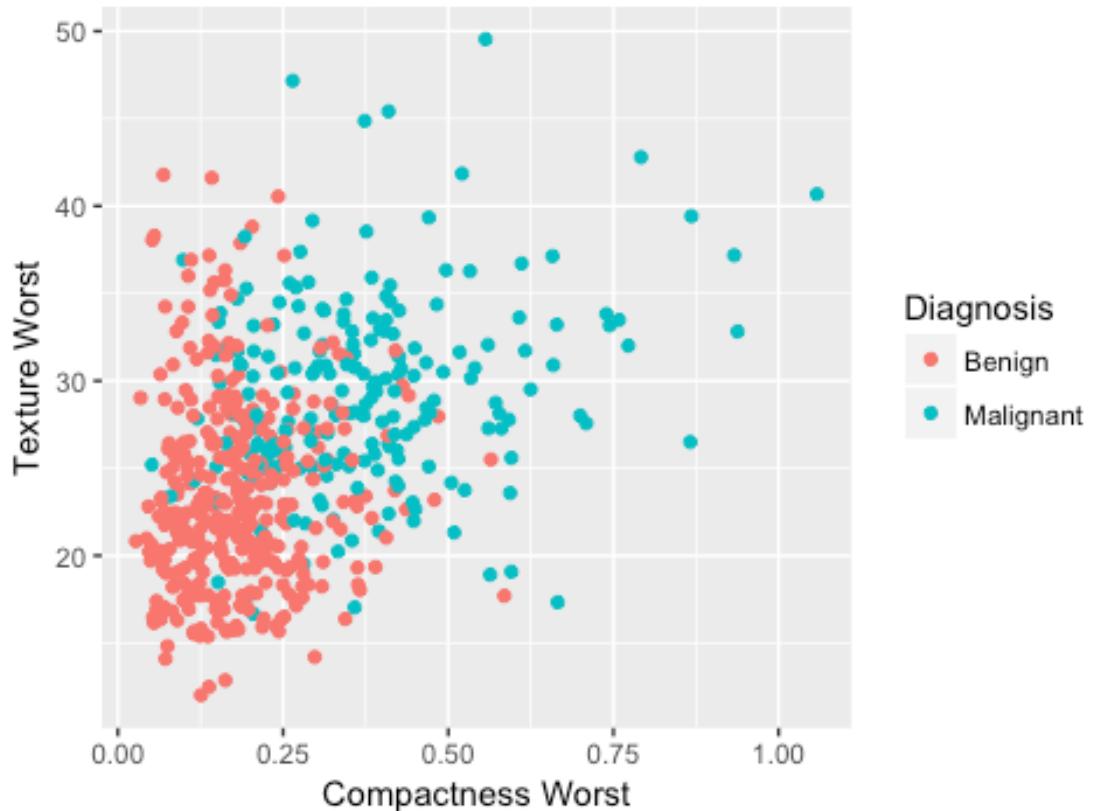
```
#radius_worst vs. compactness_worst
compactness_radius_worst_scatter <- ggplot(data = cancer, aes(x = compactness_worst, y = radius_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Worst", y = "Radius Worst", title = "Radius Worst vs. Compactness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_radius_worst_scatter
```

Radius Worst vs. Compactness Worst



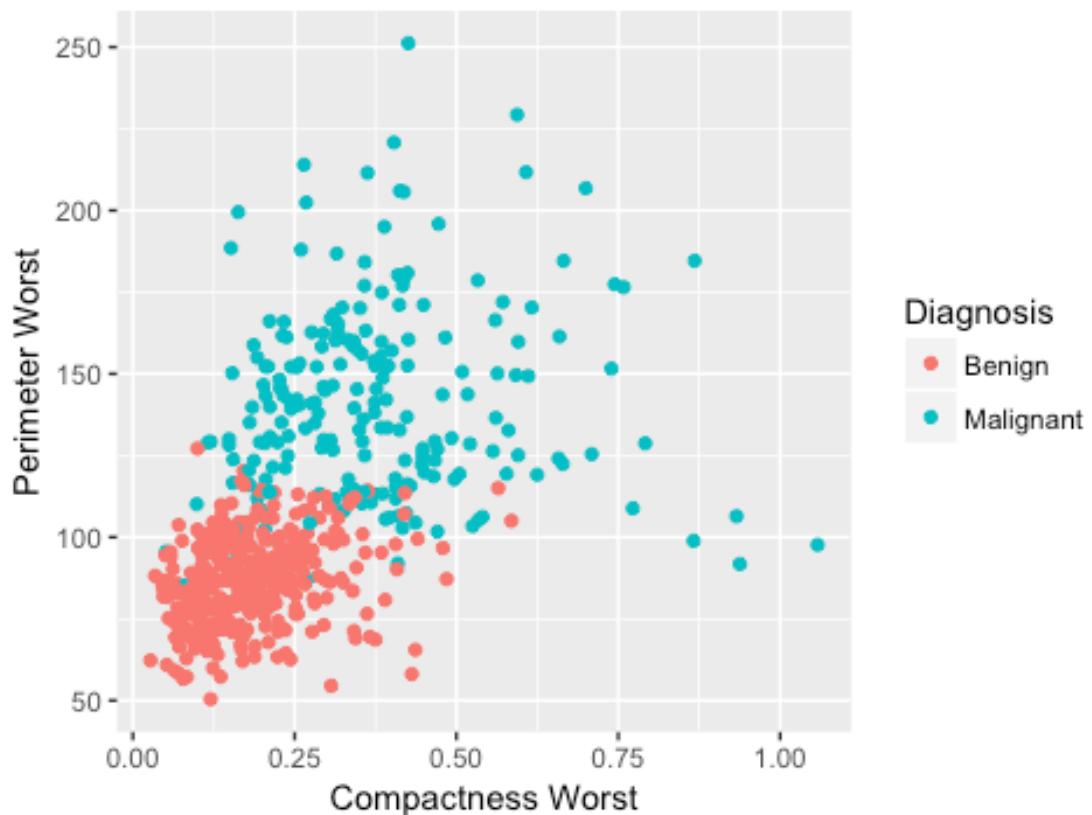
```
#texture_worst vs. compactness_worst
compactness_texture_worst_scatter <- ggplot(data = cancer, aes(x = compactness_worst, y = texture_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Worst", y = "Texture Worst", title = "Texture Worst vs. Compactness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_texture_worst_scatter
```

Texture Worst vs. Compactness Worst



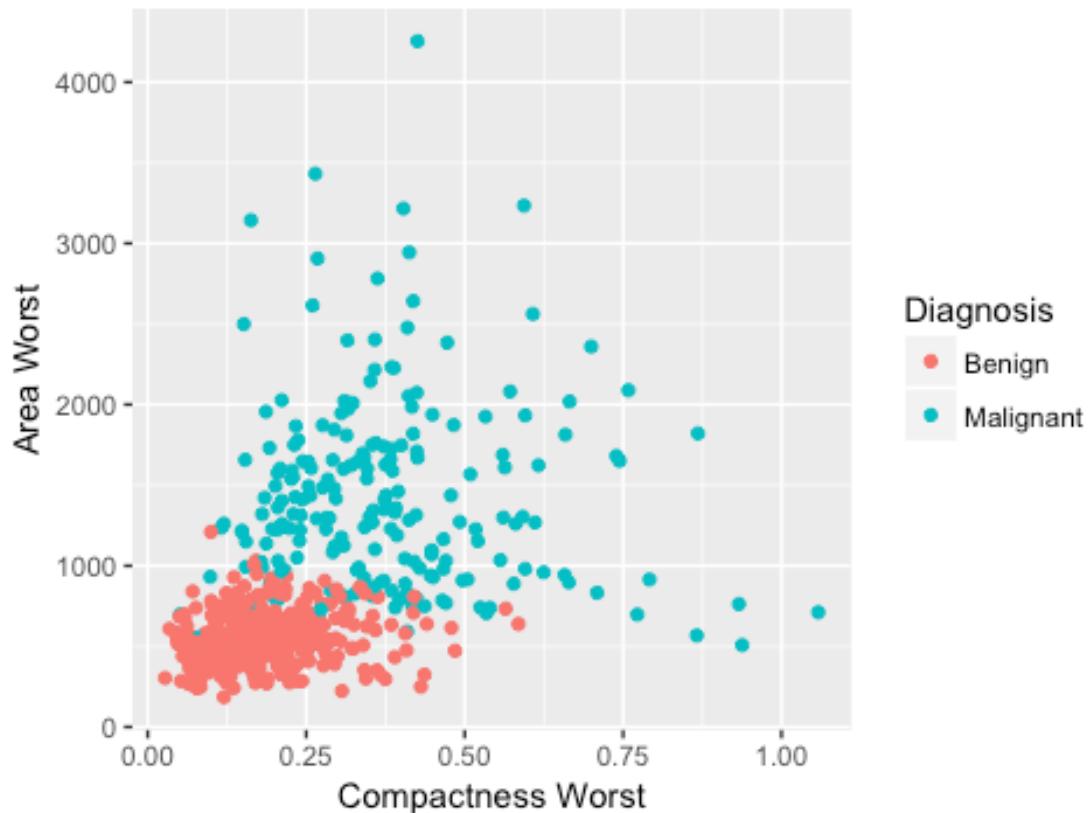
```
#perimeter_worst vs. compactness_worst
compactness_perimeter_worst_scatter <- ggplot(data = cancer, aes(x = compactness_worst, y = perimeter_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Worst", y = "Perimeter Worst", title =
  "Perimeter Worst vs. Compactness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_perimeter_worst_scatter
```

Perimeter Worst vs. Compactness Worst



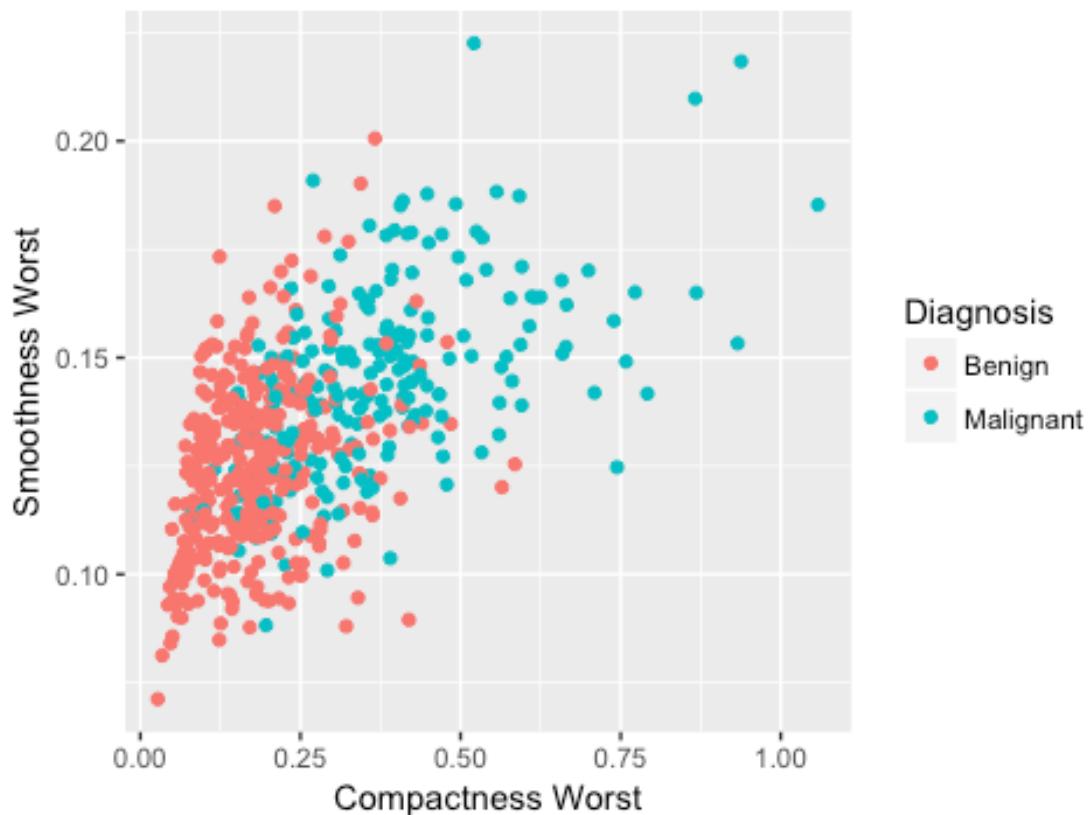
```
#area_worst vs. compactness_worst
compactness_area_worst_scatter <- ggplot(data = cancer, aes(x = compactness_worst, y = area_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Worst", y = "Area Worst", title = "Area Worst vs. Compactness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_area_worst_scatter
```

Area Worst vs. Compactness Worst



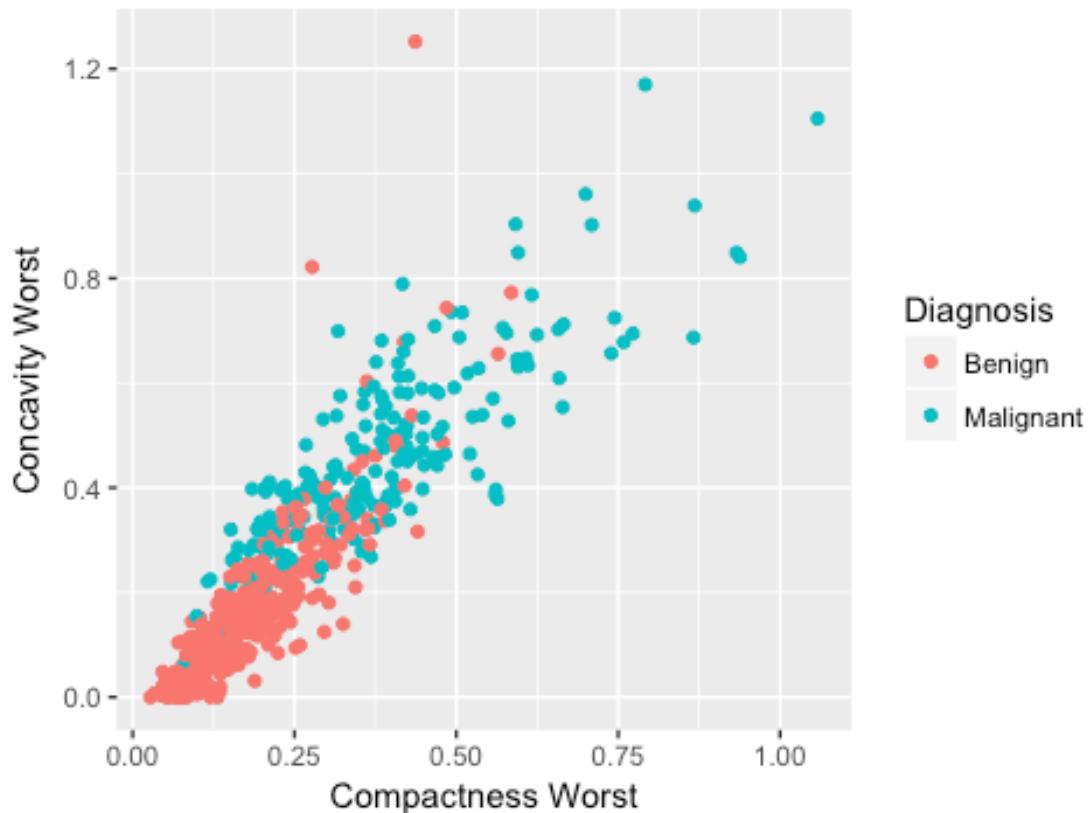
```
#smoothness_worst vs. compactness_worst
compactness_smoothness_worst_scatter <- ggplot(data = cancer, aes(x = compactness_worst, y = smoothness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Worst", y = "Smoothness Worst", title =
  "Smoothness Worst vs. Compactness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_smoothness_worst_scatter
```

Smoothness Worst vs. Compactness Worst



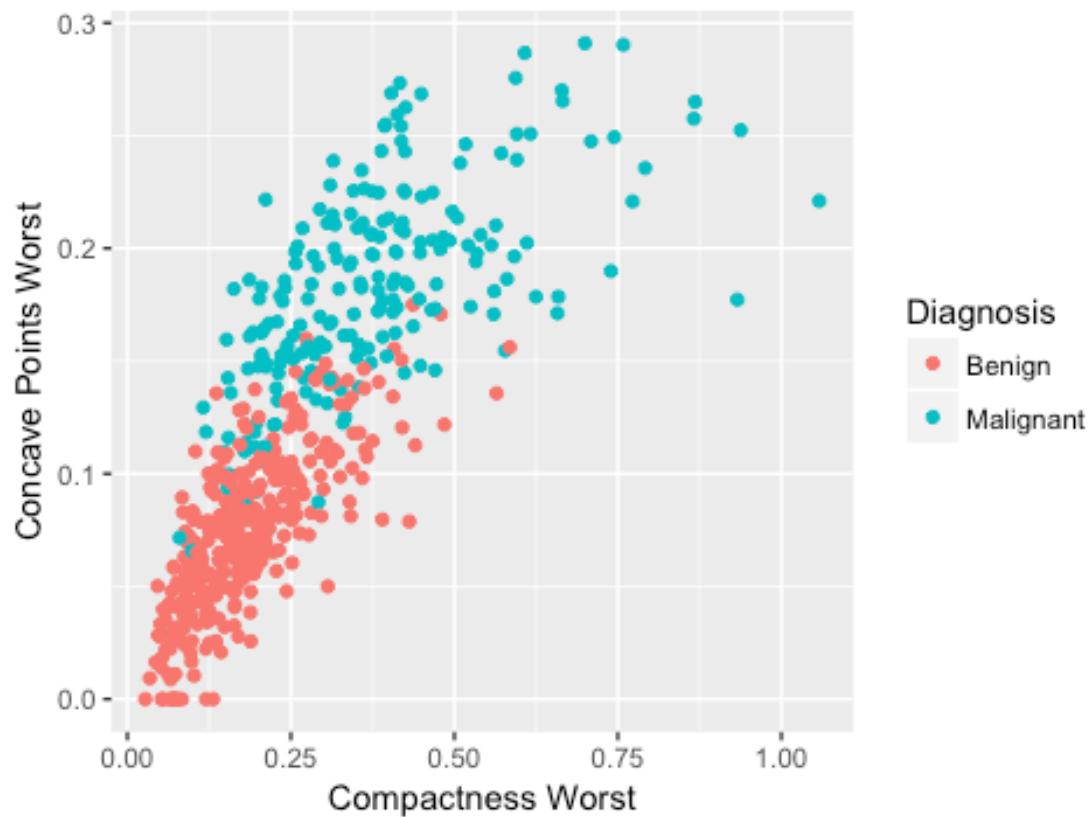
```
#concavity_worst vs. compactness_worst
compactness_concavity_worst_scatter <- ggplot(data = cancer, aes(x = compactness_worst, y = concavity_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Worst", y = "Concavity Worst", title =
  "Concavity Worst vs. Compactness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_concavity_worst_scatter
```

Concavity Worst vs. Compactness Worst



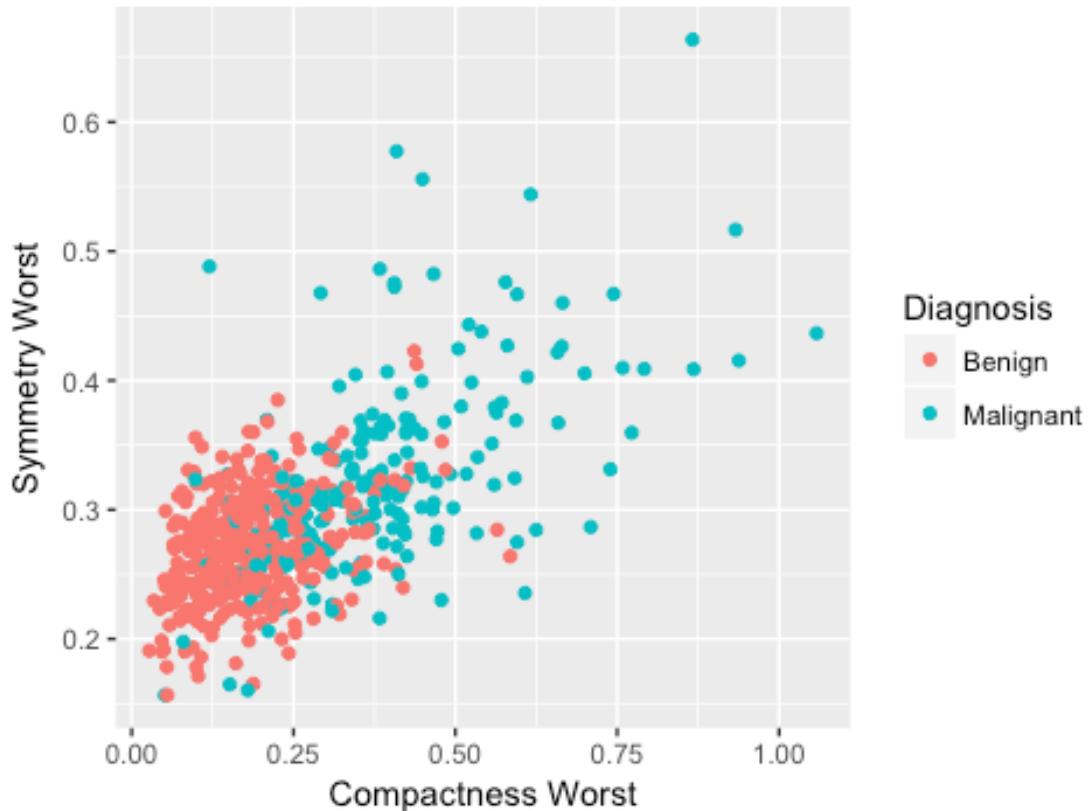
```
#concave.points_worst vs. compactness_worst
compactness_concave.points_worst_scatter <- ggplot(data = cancer, aes(x = compactness_worst, y = concave.points_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Worst", y = "Concave Points Worst", title = "Concave Points Worst vs. Compactness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_concave.points_worst_scatter
```

Concave Points Worst vs. Compactness Worst



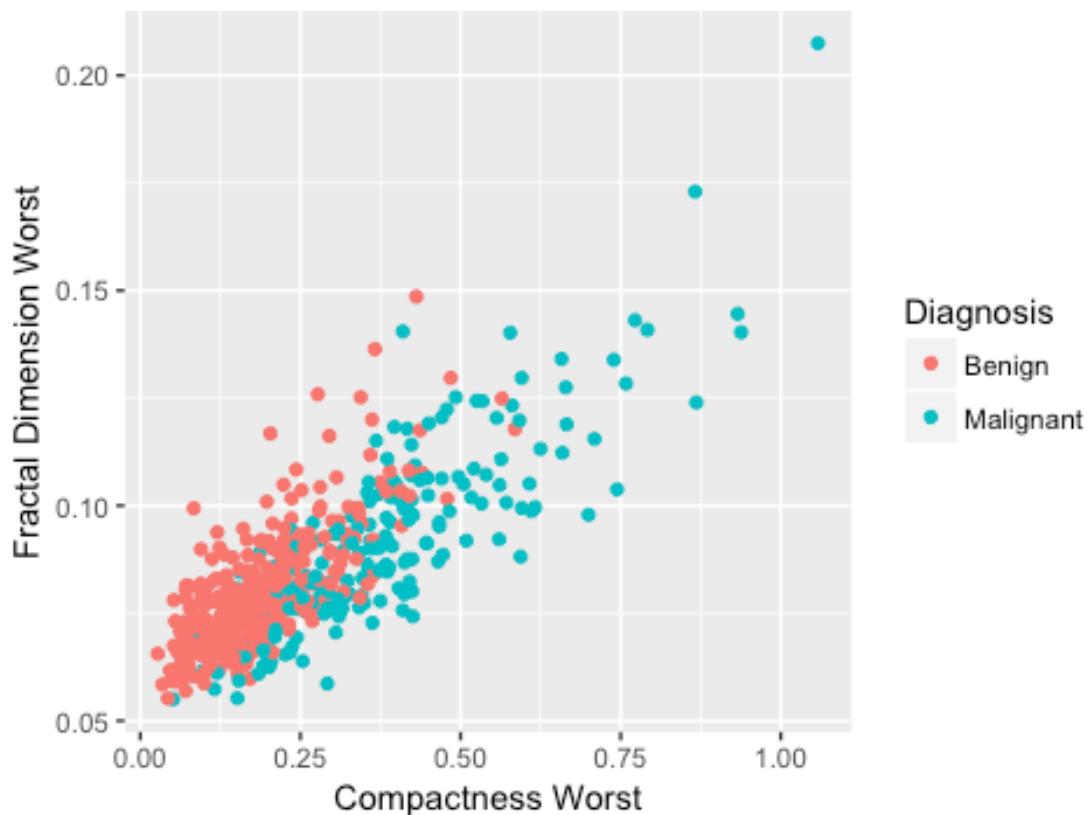
```
#symmetry_worst vs. compactness_worst
compactness_symmetry_worst_scatter <- ggplot(data = cancer, aes(x = compactness_worst, y = symmetry_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Worst", y = "Symmetry Worst", title =
  "Symmetry Worst vs. Compactness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_symmetry_worst_scatter
```

Symmetry Worst vs. Compactness Worst



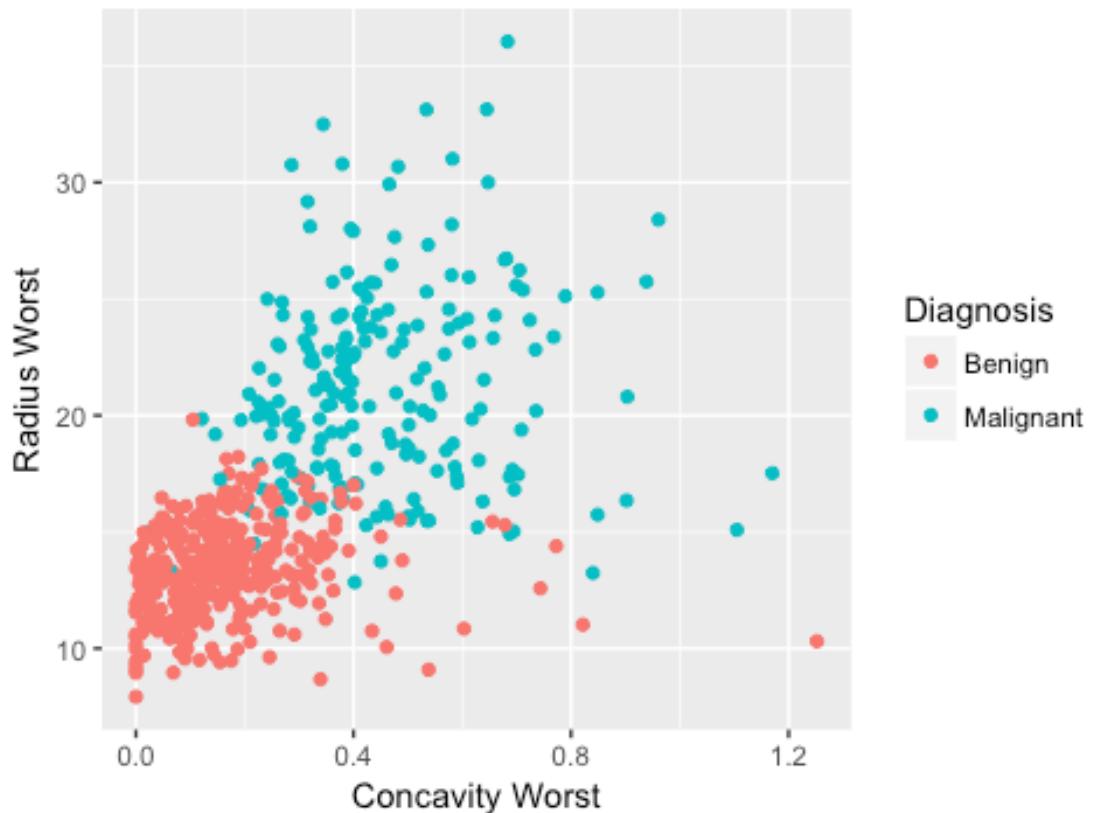
```
#fractal_dimension_worst vs. compactness_worst
compactness_fractal_dimension_worst_scatter <- ggplot(data = cancer, aes(x = compactness_worst, y = fractal_dimension_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Worst", y = "Fractal Dimension Worst",
  title = "Fractal Dimension Worst vs. Compactness Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_fractal_dimension_worst_scatter
```

Fractal Dimension Worst vs. Compactness Worst



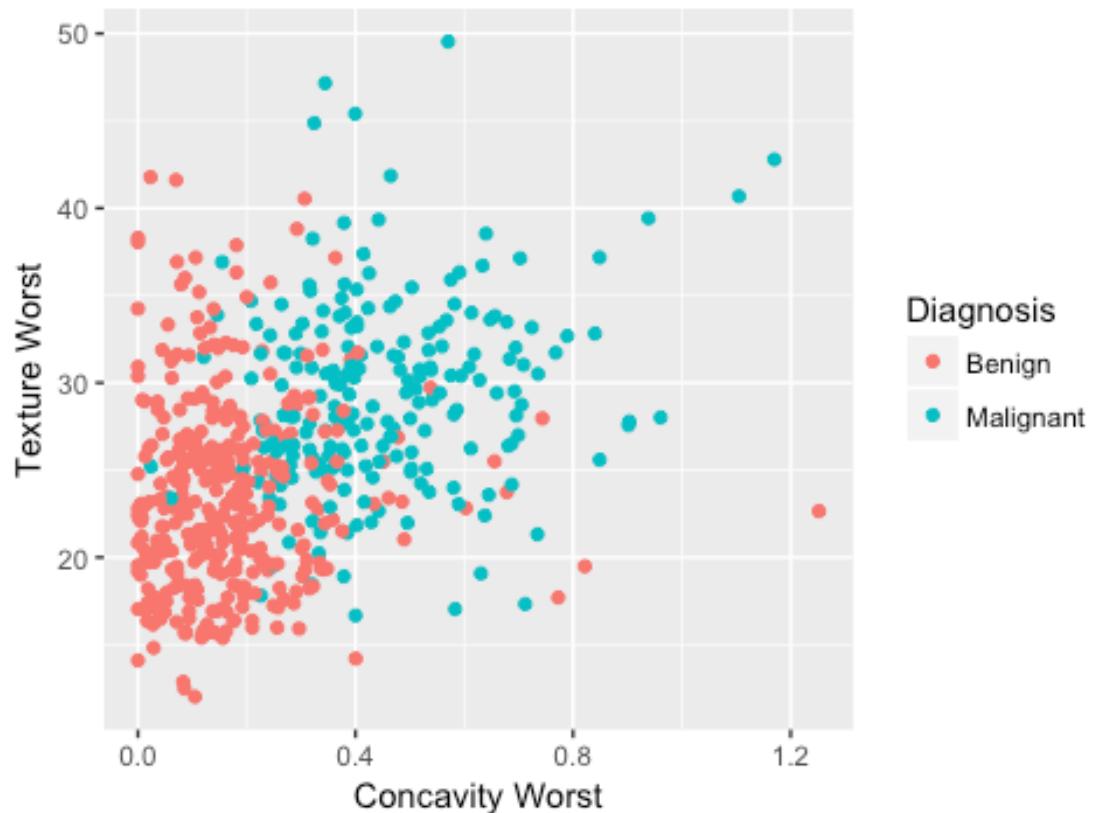
```
#radius_worst vs. concavity_worst
concavity_radius_worst_scatter <- ggplot(data = cancer, aes(x = concavity_worst, y = radius_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Worst", y = "Radius Worst", title = "Radius Worst vs. Concavity Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_radius_worst_scatter
```

Radius Worst vs. Concavity Worst



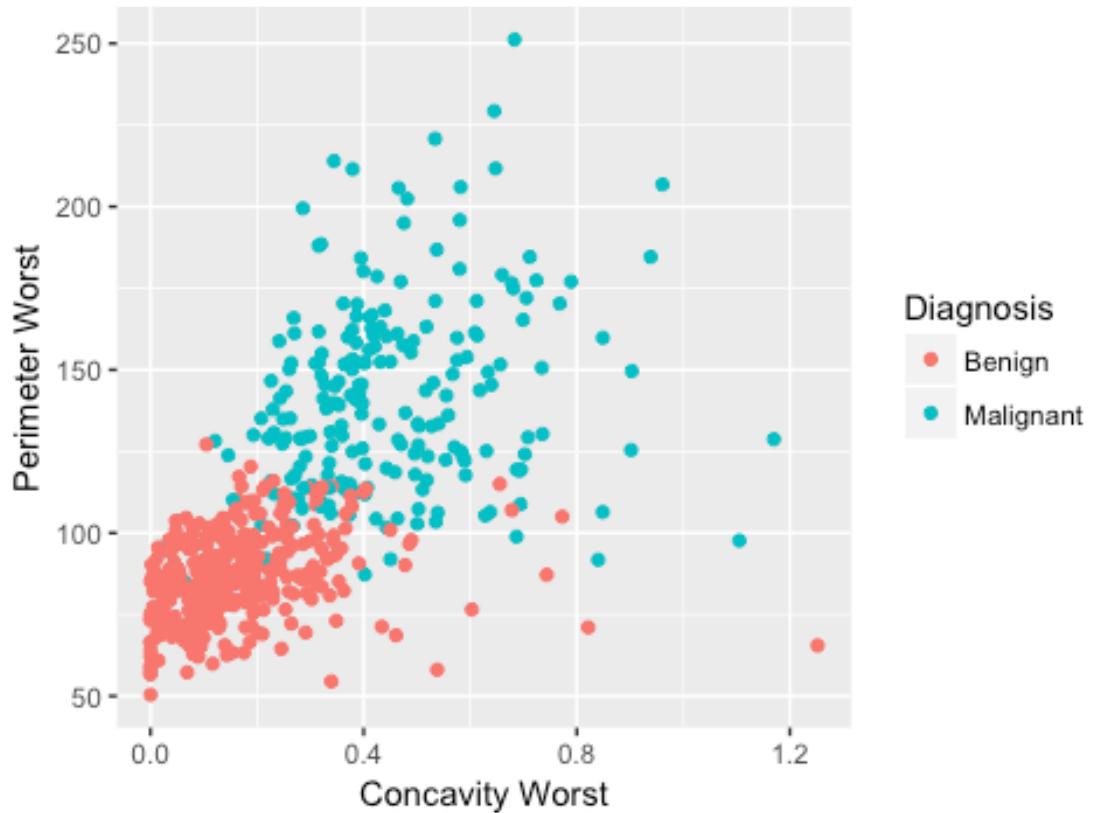
```
#texture_worst vs. concavity_worst
concavity_texture_worst_scatter <- ggplot(data = cancer, aes(x = concavity_worst, y = texture_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Worst", y = "Texture Worst", title = "Texture Worst vs. Concavity Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_texture_worst_scatter
```

Texture Worst vs. Concavity Worst



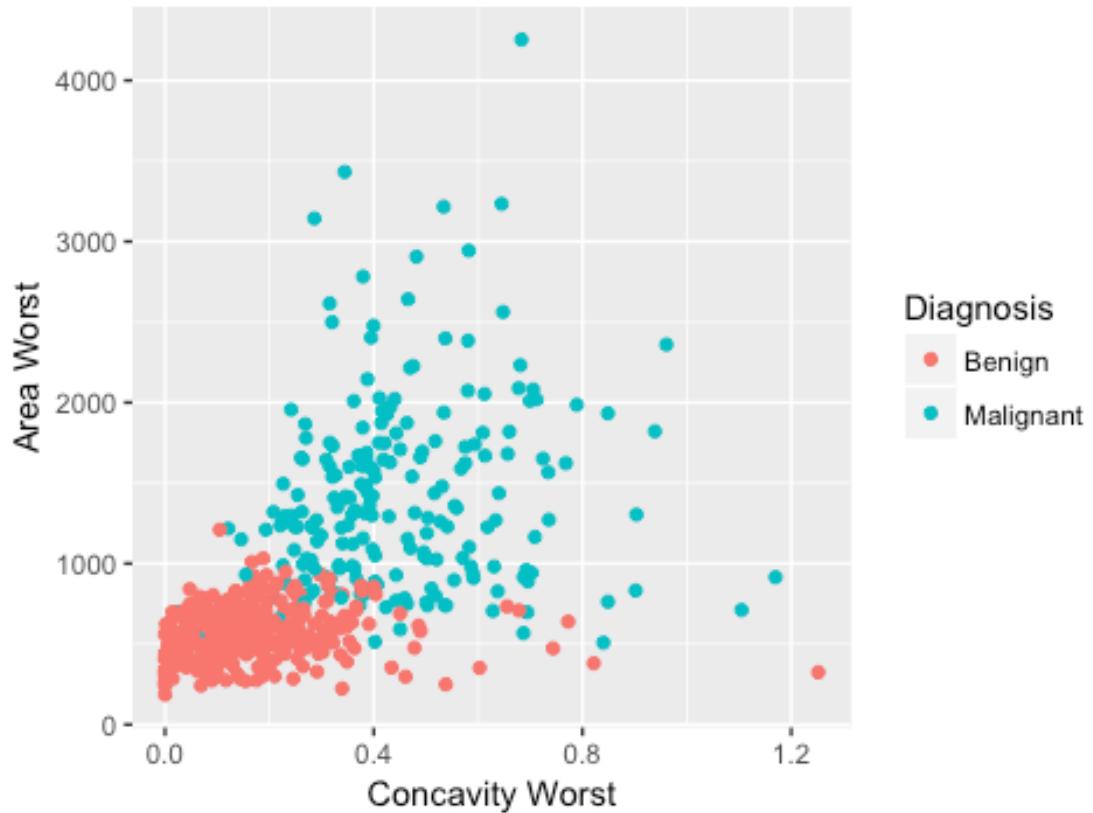
```
#perimeter_worst vs. concavity_worst
concavity_perimeter_worst_scatter <- ggplot(data = cancer, aes(x = concavity_worst, y = perimeter_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Worst", y = "Perimeter Worst", title = "Perimeter Worst vs. Concavity Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_perimeter_worst_scatter
```

Perimeter Worst vs. Concavity Worst



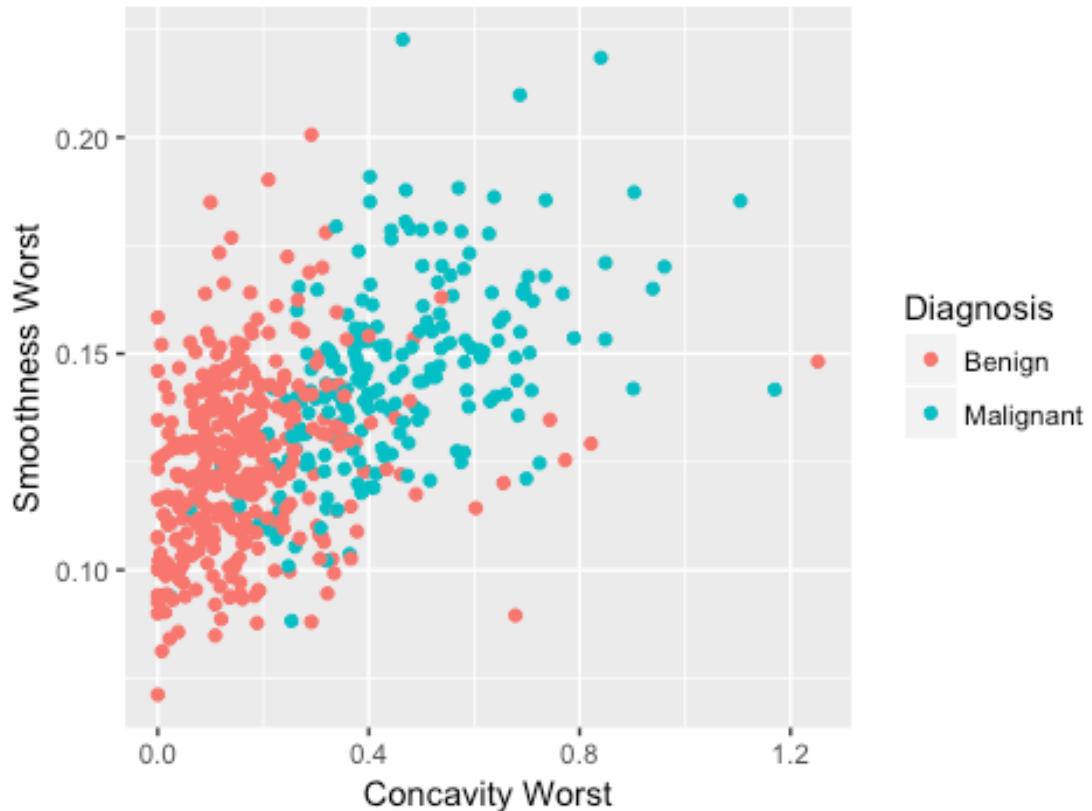
```
#area_worst vs. concavity_worst
concavity_area_worst_scatter <- ggplot(data = cancer, aes(x = concavity_worst,
, y = area_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Worst", y = "Area Worst", title = "Area
Worst vs. Concavity Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_area_worst_scatter
```

Area Worst vs. Concavity Worst



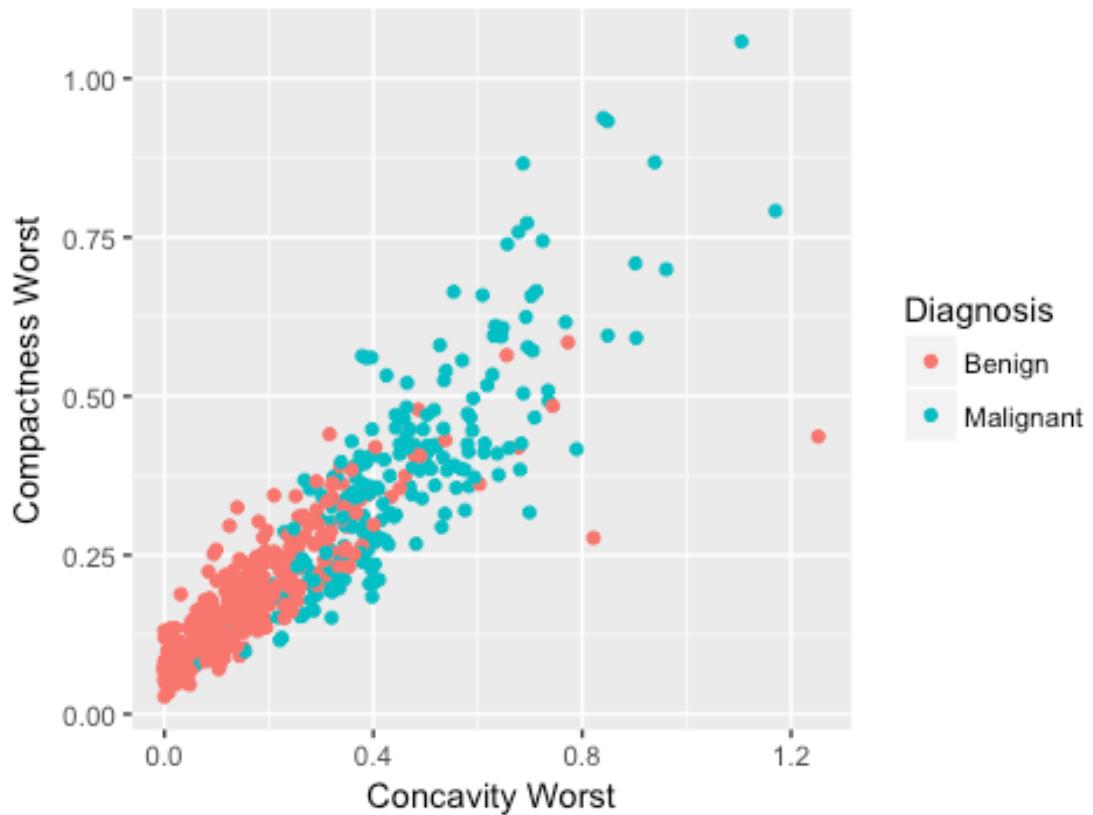
```
#smoothness_worst vs. concavity_worst
concavity_smoothness_worst_scatter <- ggplot(data = cancer, aes(x = concavity_worst, y = smoothness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Worst", y = "Smoothness Worst", title =
"Smoothness Worst vs. Concavity Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_smoothness_worst_scatter
```

Smoothness Worst vs. Concavity Worst



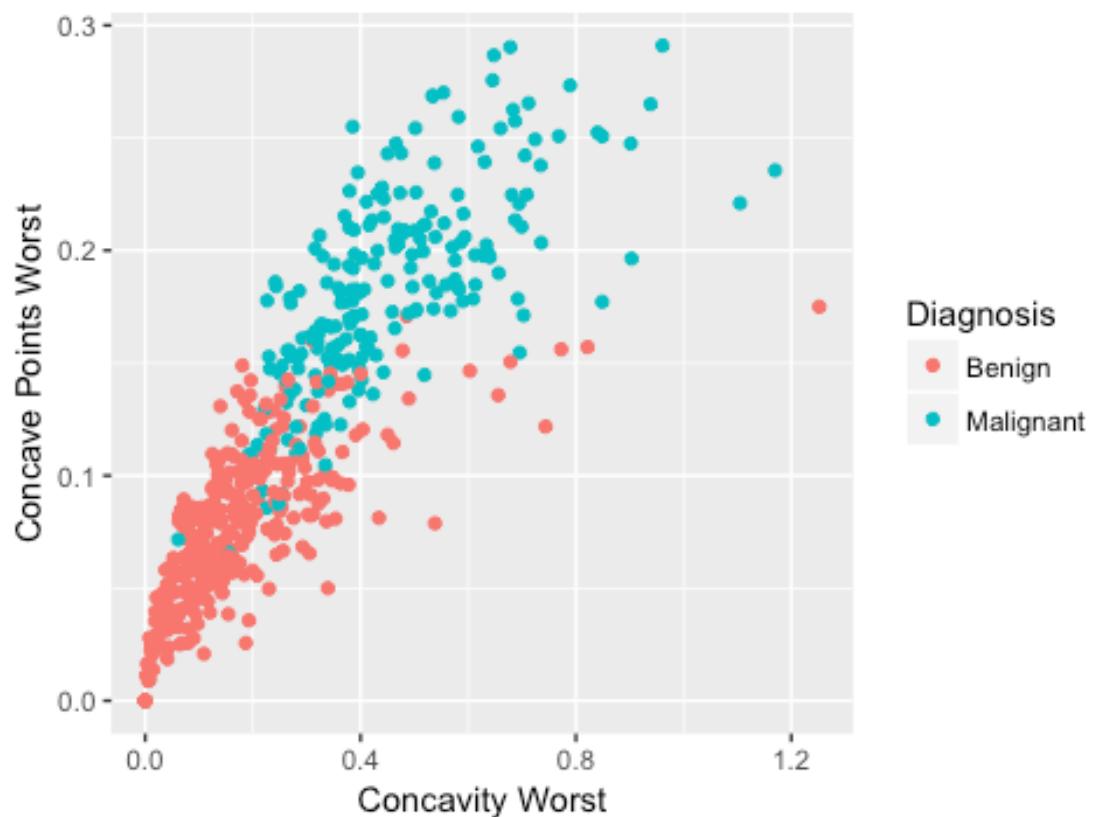
```
#compactness_worst vs. concavity_worst
concavity_compactness_worst_scatter <- ggplot(data = cancer, aes(x = concavity_worst, y = compactness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Worst", y = "Compactness Worst", title = "Compactness Worst vs. Concavity Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_compactness_worst_scatter
```

Compactness Worst vs. Concavity Worst



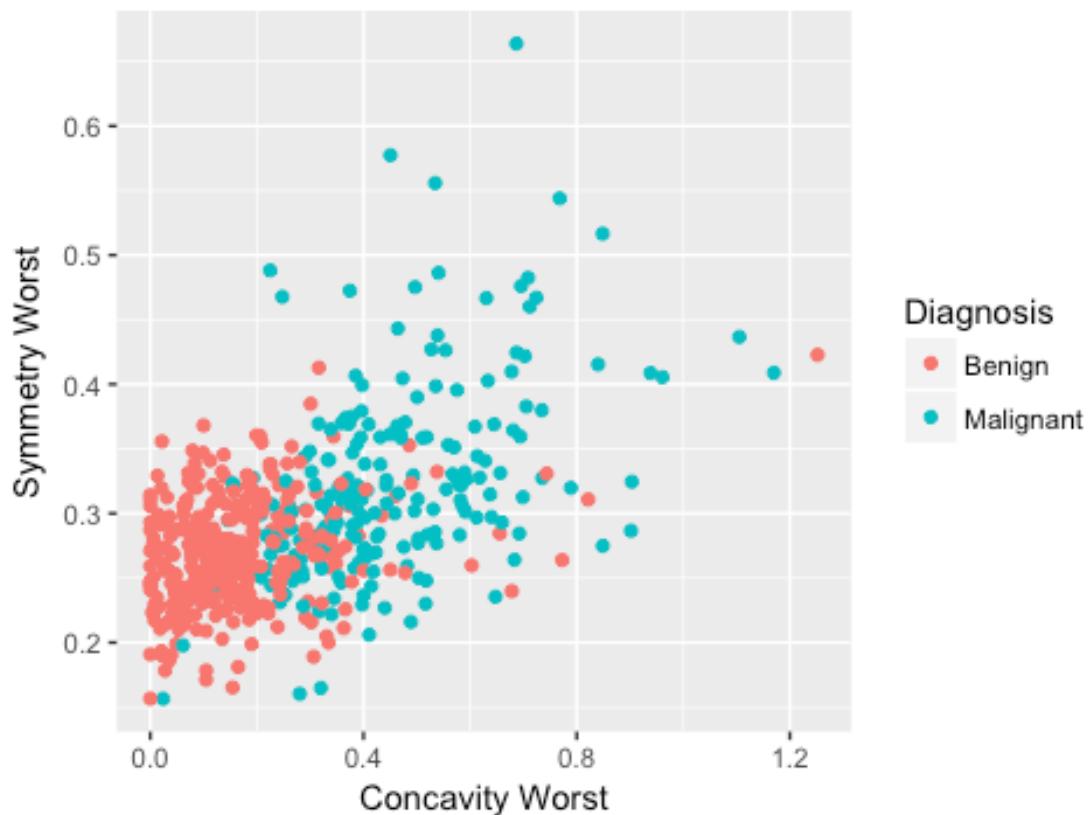
```
#concave.points_worst vs. concavity_worst
concavity_concave.points_worst_scatter <- ggplot(data = cancer, aes(x = concavity_worst, y = concave.points_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Worst", y = "Concave Points Worst", title = "Concave Points Worst vs. Concavity Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_concave.points_worst_scatter
```

Concave Points Worst vs. Concavity Worst



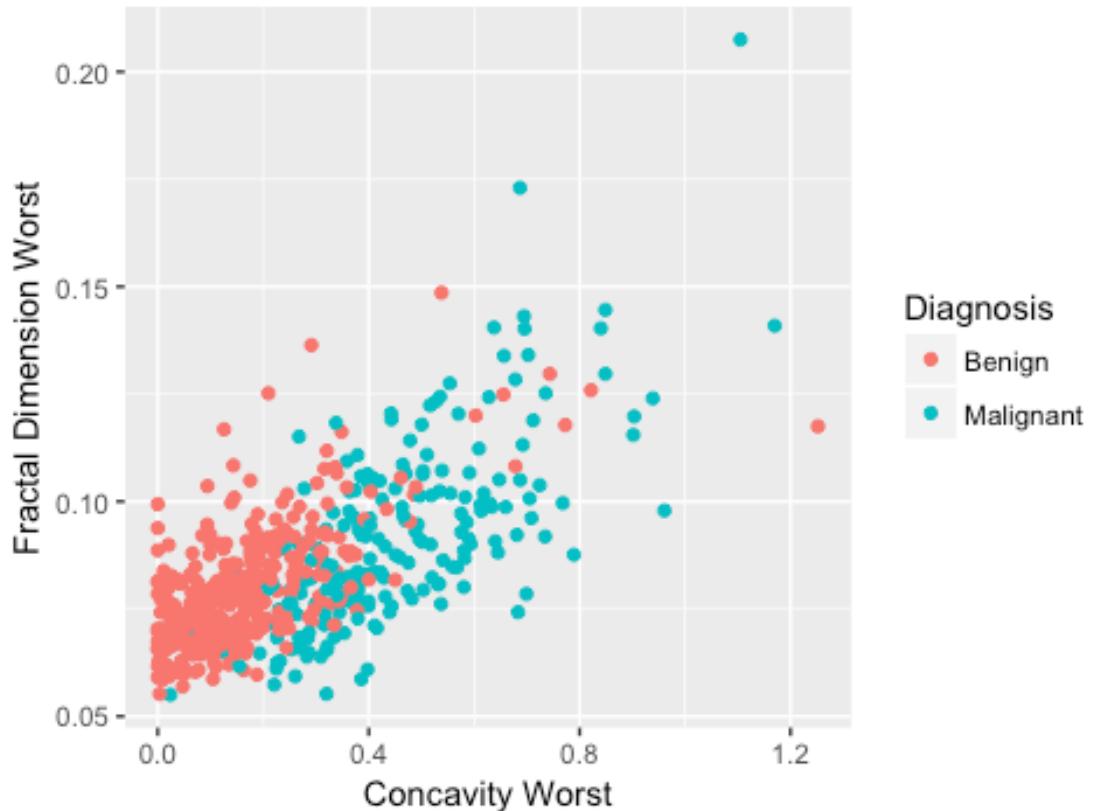
```
#symmetry_worst vs. concavity_worst
concavity_symmetry_worst_scatter <- ggplot(data = cancer, aes(x = concavity_worst, y = symmetry_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Worst", y = "Symmetry Worst", title = "Symmetry Worst vs. Concavity Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_symmetry_worst_scatter
```

Symmetry Worst vs. Concavity Worst



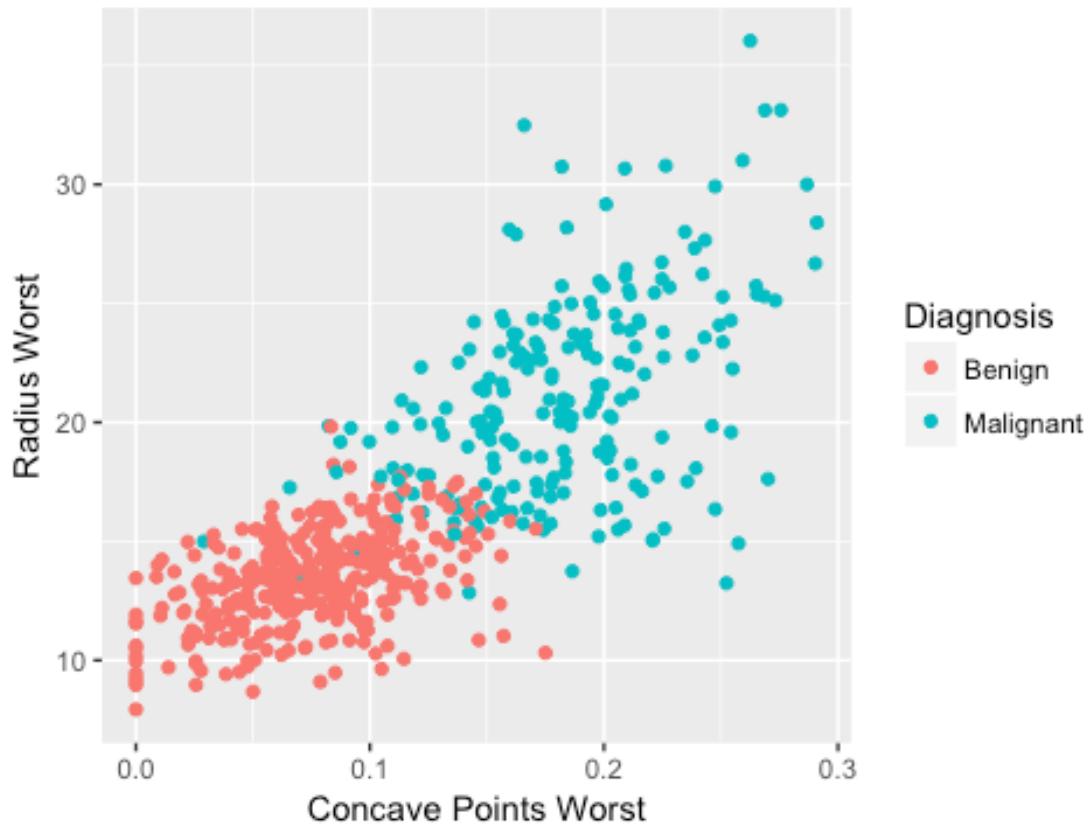
```
#fractal_dimension_worst vs. concavity_worst
concavity_fractal_dimension_worst_scatter <- ggplot(data = cancer, aes(x = concavity_worst, y = fractal_dimension_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Worst", y = "Fractal Dimension Worst", title = "Fractal Dimension Worst vs. Concavity Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_fractal_dimension_worst_scatter
```

Fractal Dimension Worst vs. Concavity Worst



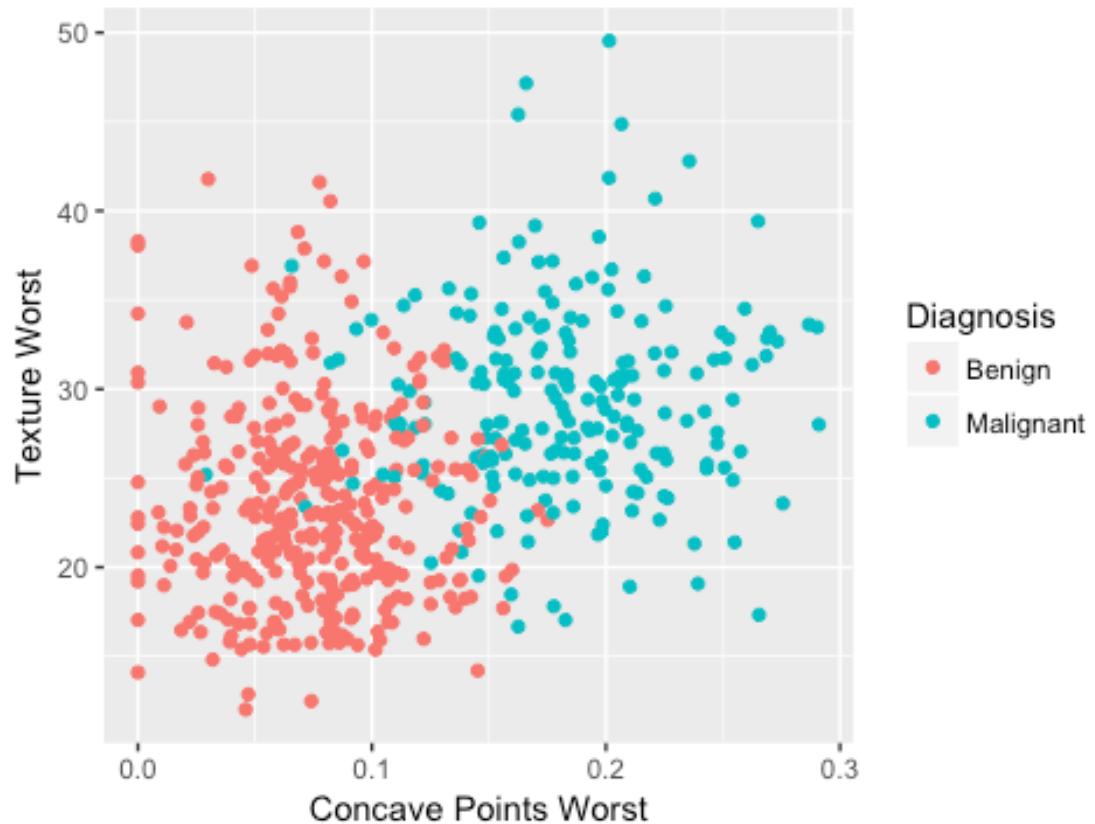
```
#radius_worst vs. concave.points_worst
concave.points_radius_worst_scatter <- ggplot(data = cancer, aes(x = concave.
points_worst, y = radius_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Worst", y = "Radius Worst", title =
"Radius Worst vs. Concave Points Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_radius_worst_scatter
```

Radius Worst vs. Concave Points Worst



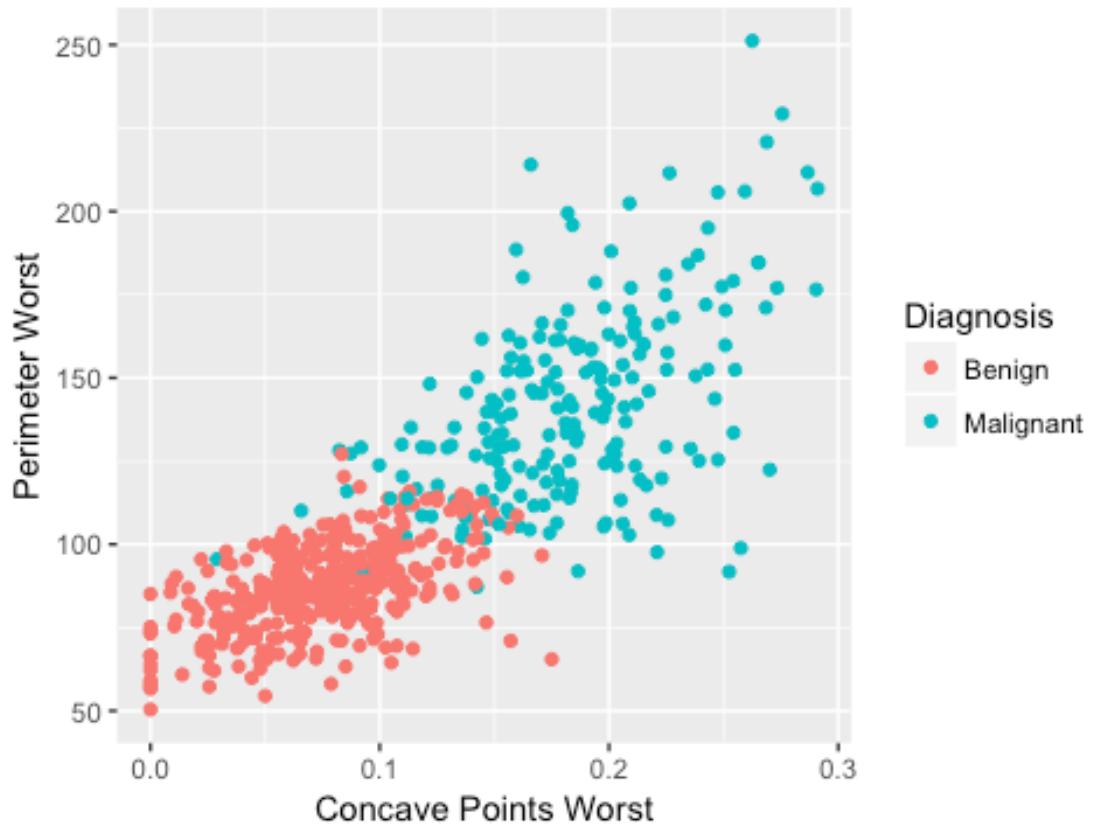
```
#texture_worst vs. concave.points_worst
concave.points_texture_worst_scatter <- ggplot(data = cancer, aes(x = concave
  .points_worst, y = texture_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Worst", y = "Texture Worst", title
  = "Texture Worst vs. Concave Points Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_texture_worst_scatter
```

Texture Worst vs. Concave Points Worst



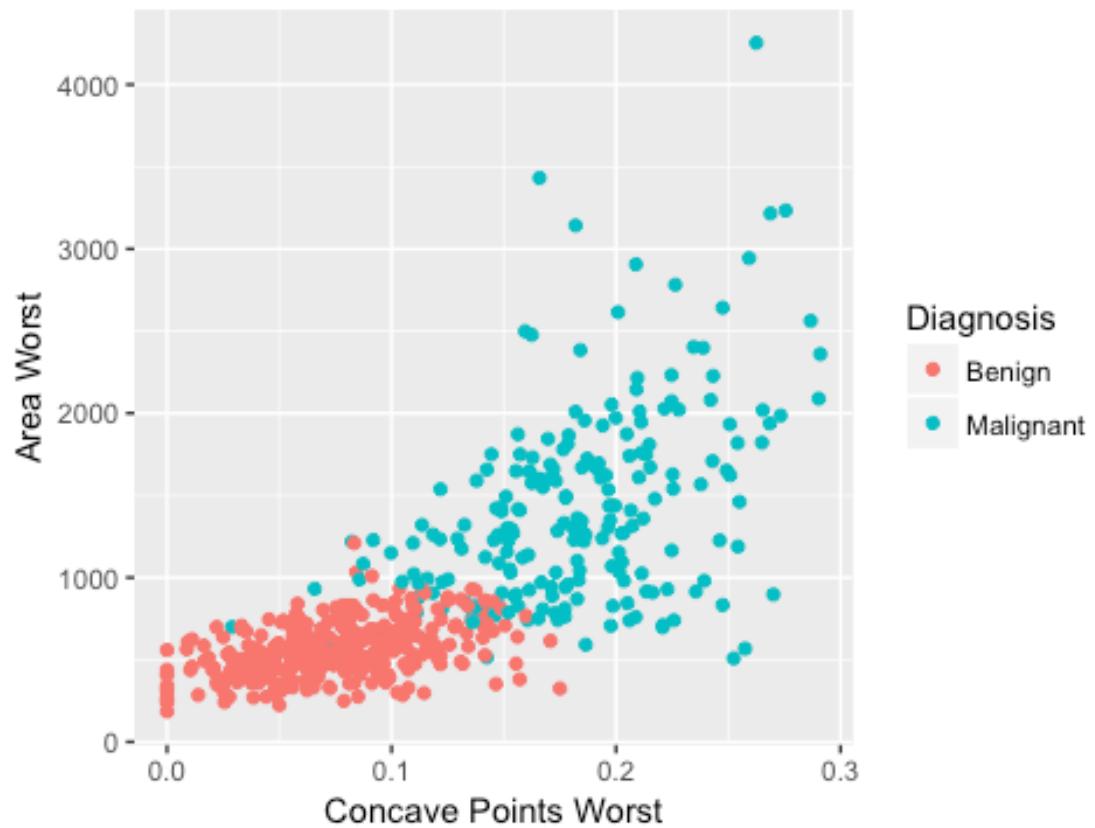
```
#perimeter_worst vs. concave.points_worst
concave.points_perimeter_worst_scatter <- ggplot(data = cancer, aes(x = concave.points_worst, y = perimeter_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Worst", y = "Perimeter Worst", title = "Perimeter Worst vs. Concave Points Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_perimeter_worst_scatter
```

Perimeter Worst vs. Concave Points Worst



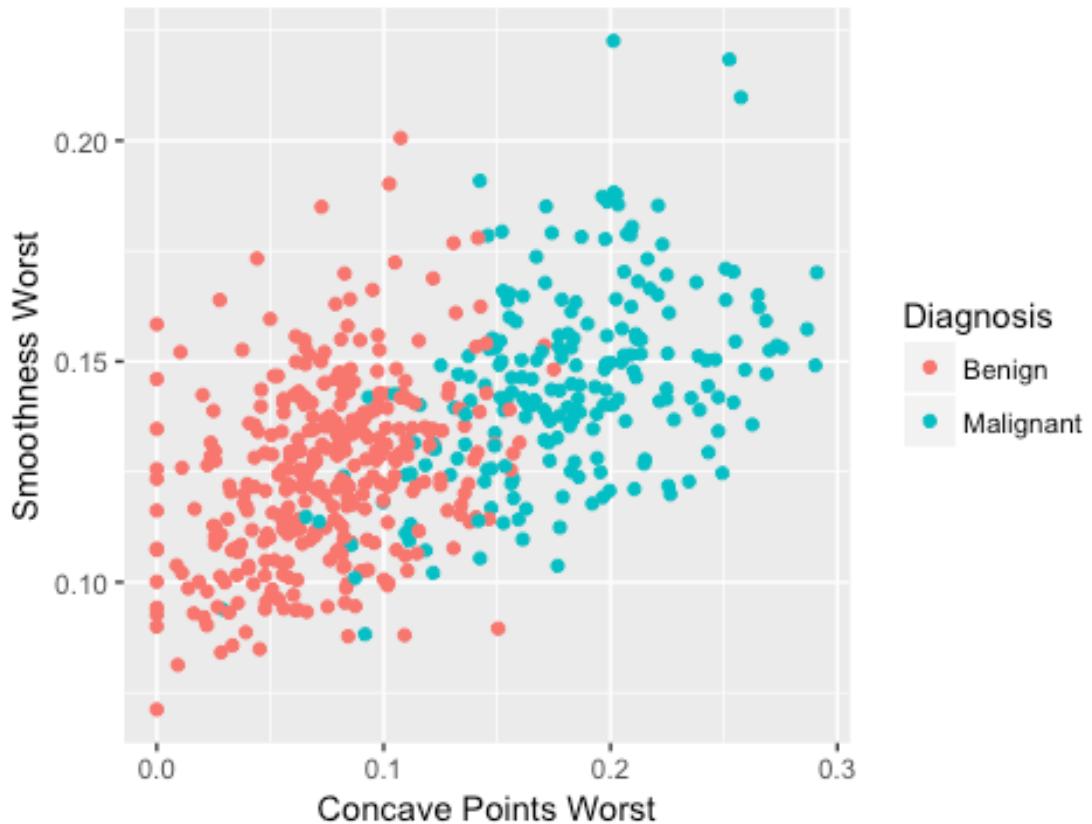
```
#area_worst vs. concave.points_worst
concave.points_area_worst_scatter <- ggplot(data = cancer, aes(x = concave.points_worst, y = area_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Worst", y = "Area Worst", title = "Area Worst vs. Concave Points Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_area_worst_scatter
```

Area Worst vs. Concave Points Worst



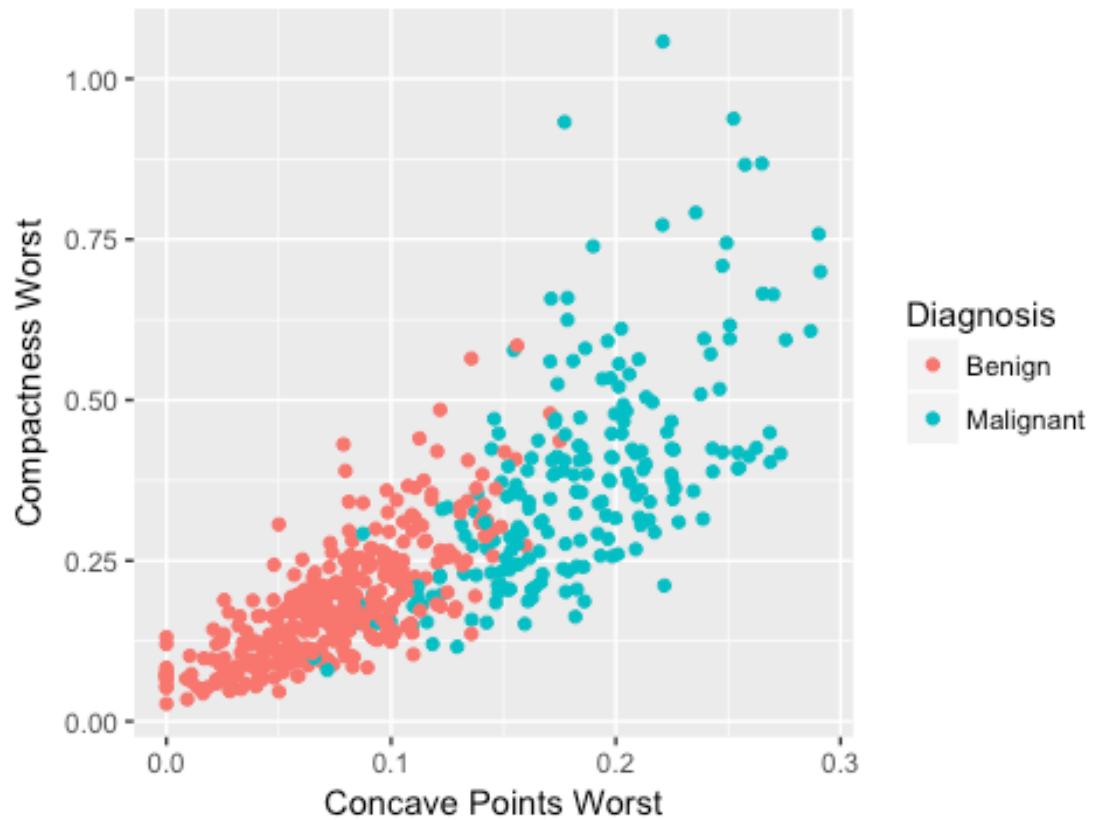
```
#smoothness_worst vs. concave.points_worst
concave.points_smoothness_worst_scatter <- ggplot(data = cancer, aes(x = concave.points_worst, y = smoothness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Worst", y = "Smoothness Worst", title = "Smoothness Worst vs. Concave Points Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_smoothness_worst_scatter
```

Smoothness Worst vs. Concave Points Worst



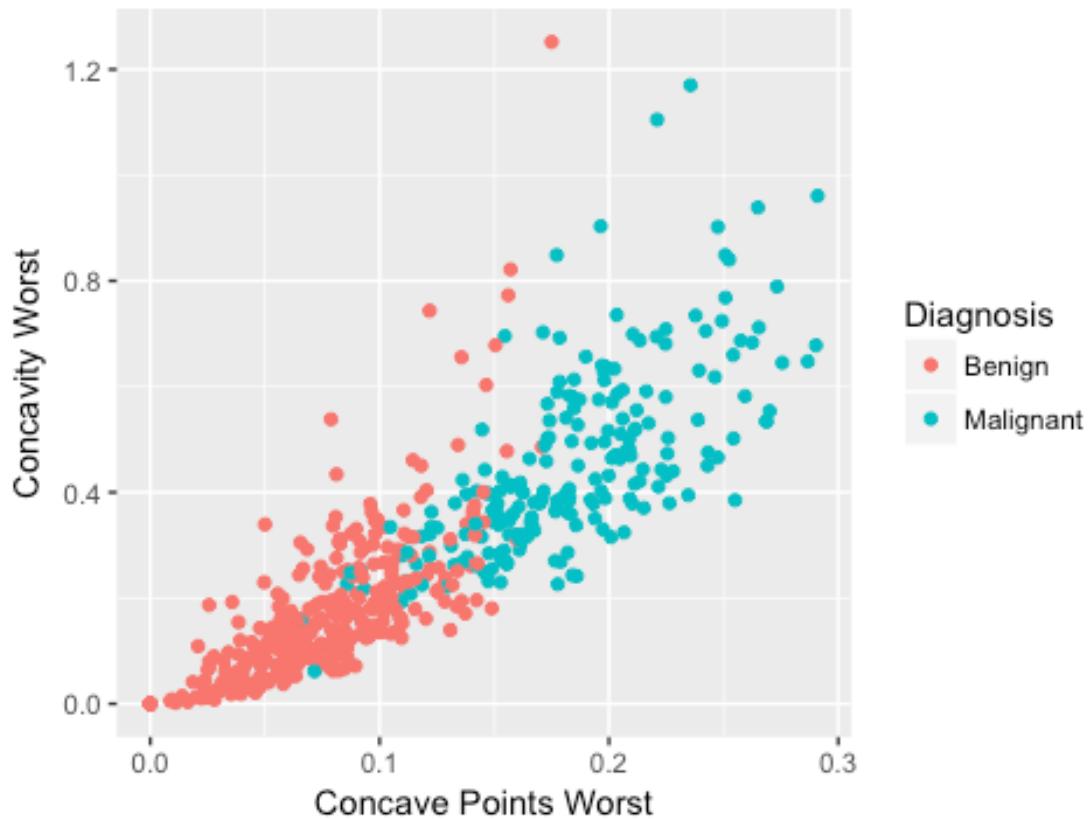
```
#compactness_worst vs. concave.points_worst
concave.points_compactness_worst_scatter <- ggplot(data = cancer, aes(x = concave.points_worst, y = compactness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Worst", y = "Compactness Worst", title = "Compactness Worst vs. Concave Points Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_compactness_worst_scatter
```

Compactness Worst vs. Concave Points Worst



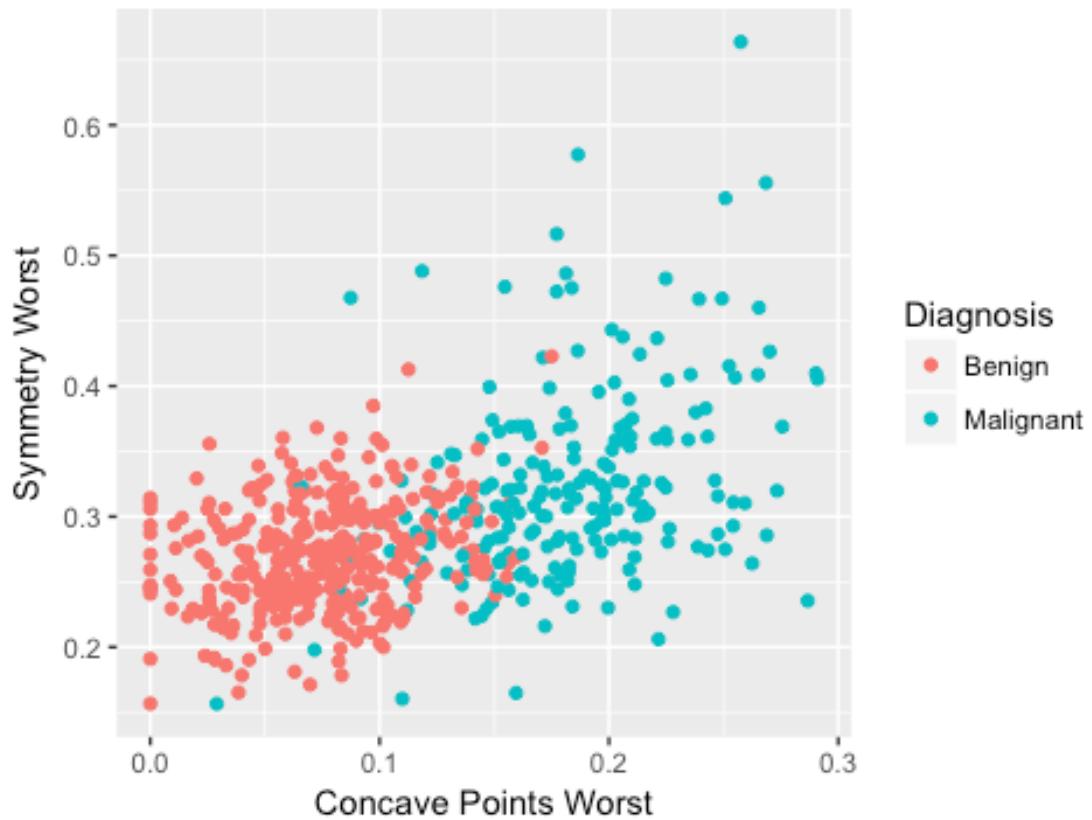
```
#concavity_worst vs. concave.points_worst
concave.points_concavity_worst_scatter <- ggplot(data = cancer, aes(x = concave.points_worst, y = concavity_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Worst", y = "Concavity Worst", title = "Concavity Worst vs. Concave Points Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_concavity_worst_scatter
```

Concavity Worst vs. Concave Points Worst



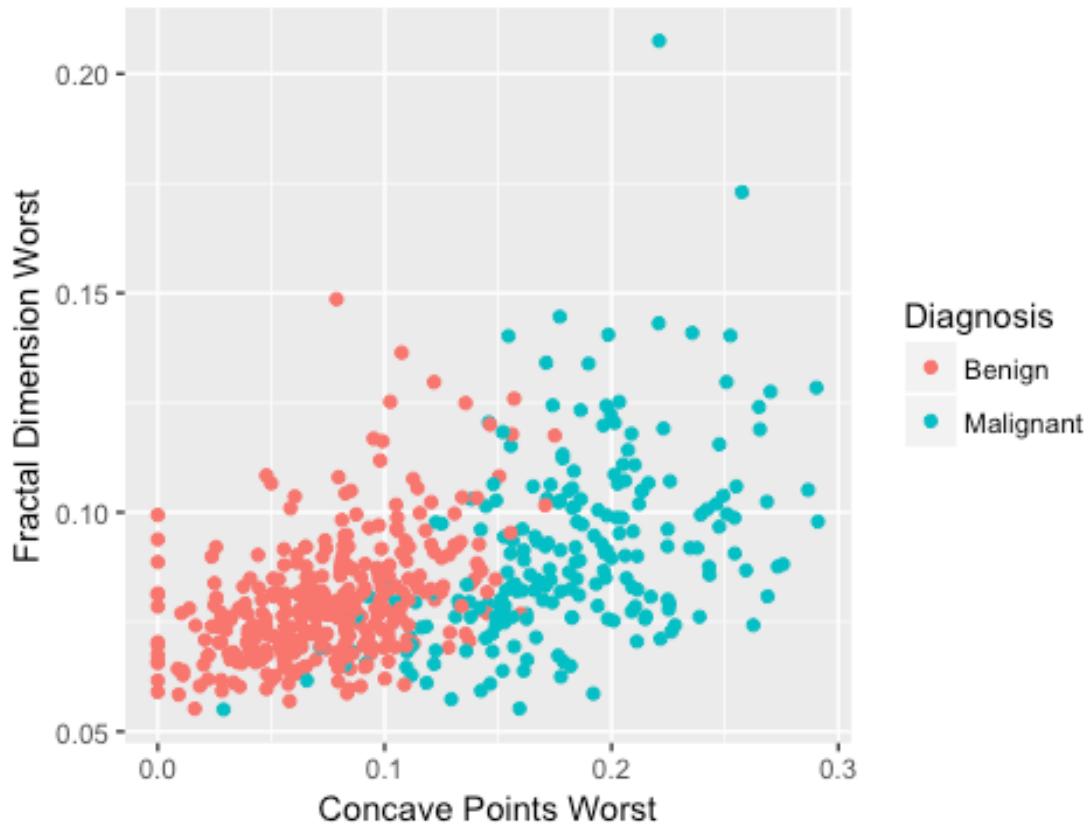
```
#symmetry_worst vs. concave.points_worst
concave.points_symmetry_worst_scatter <- ggplot(data = cancer, aes(x = concave.points_worst, y = symmetry_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Worst", y = "Symmetry Worst", title =
  "Symmetry Worst vs. Concave Points Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_symmetry_worst_scatter
```

Symmetry Worst vs. Concave Points Worst



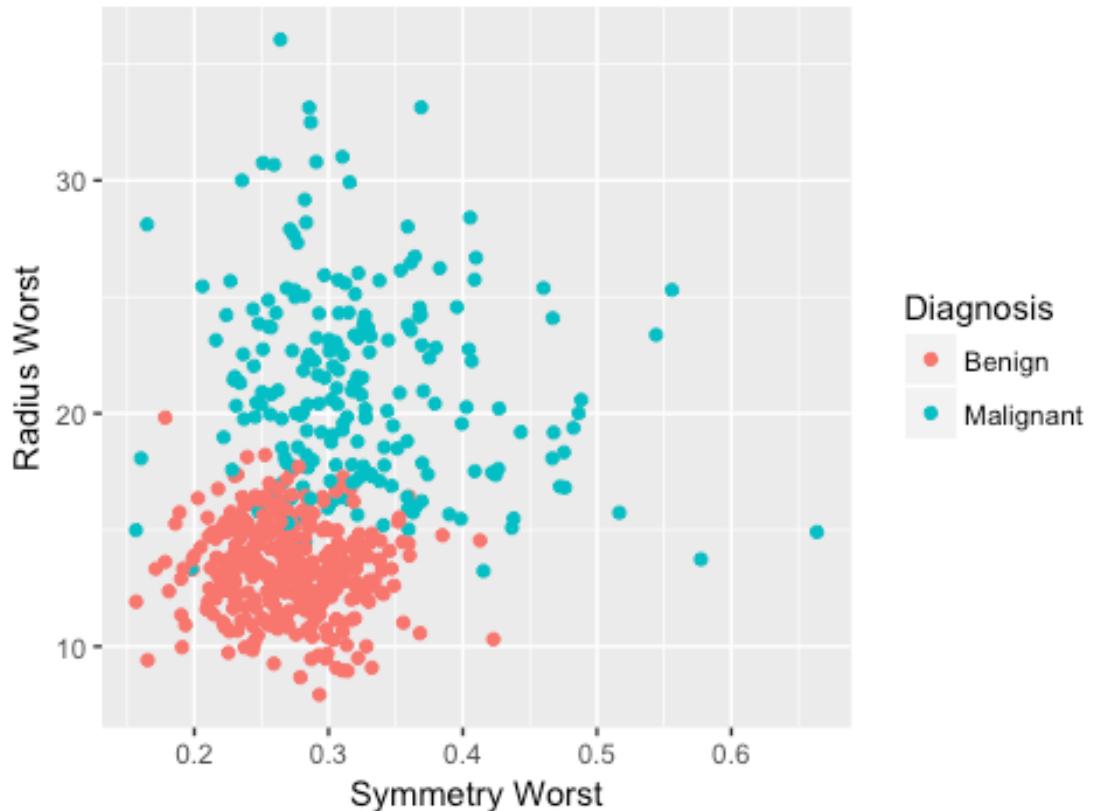
```
#fractal_dimension_worst vs. concave.points_worst
concave.points_fractal_dimension_worst_scatter <- ggplot(data = cancer, aes(x = concave.points_worst, y = fractal_dimension_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Worst", y = "Fractal Dimension Worst", title = "Fractal Dimension Worst vs. Concave Points Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_fractal_dimension_worst_scatter
```

Fractal Dimension Worst vs. Concave Points Worst



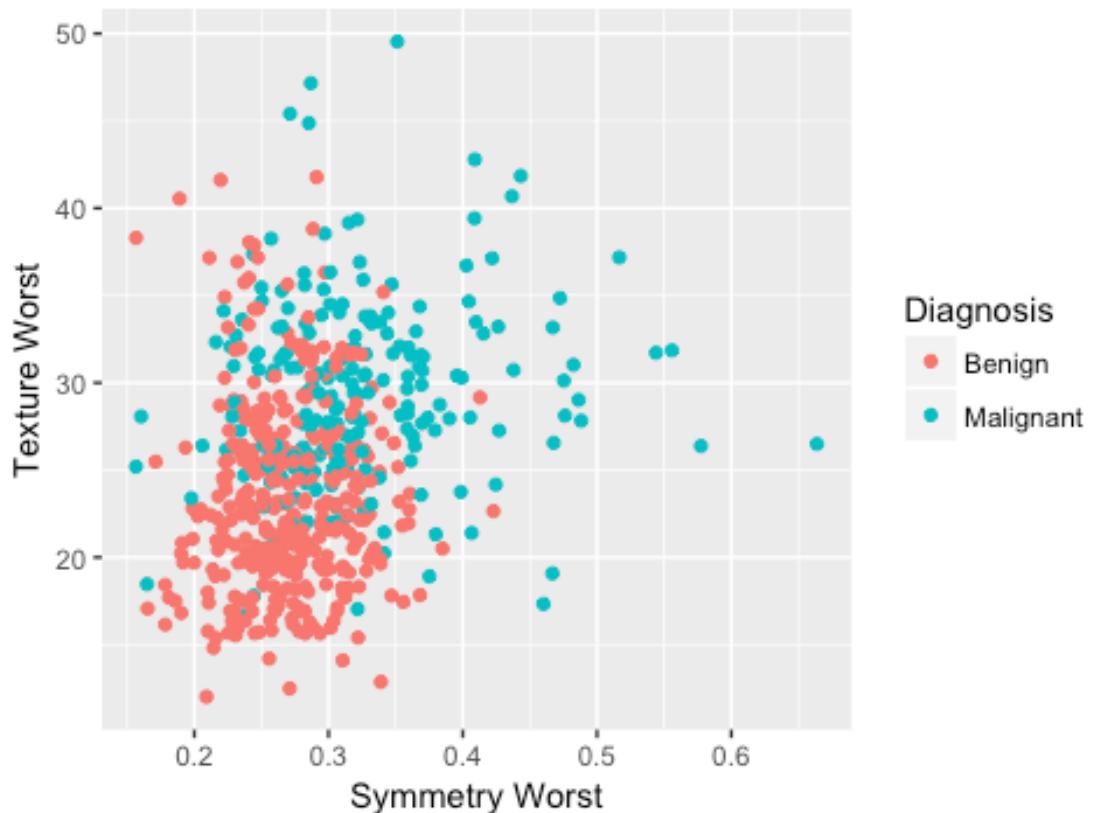
```
#radius_worst vs. symmetry_worst
symmetry_radius_worst_scatter <- ggplot(data = cancer, aes(x = symmetry_worst,
, y = radius_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Worst", y = "Radius Worst", title = "Radius Worst vs. Symmetry Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_radius_worst_scatter
```

Radius Worst vs. Symmetry Worst



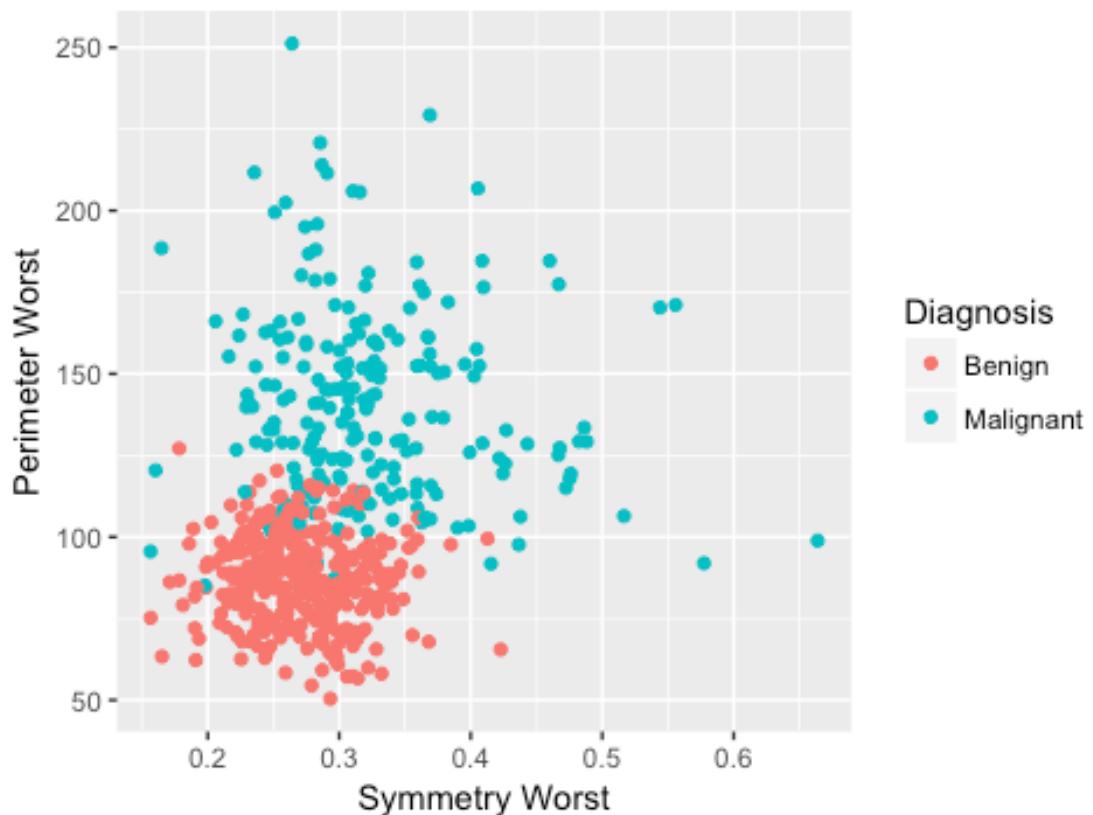
```
#texture_worst vs. symmetry_worst
symmetry_texture_worst_scatter <- ggplot(data = cancer, aes(x = symmetry_worst, y = texture_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Worst", y = "Texture Worst", title = "Texture Worst vs. Symmetry Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_texture_worst_scatter
```

Texture Worst vs. Symmetry Worst



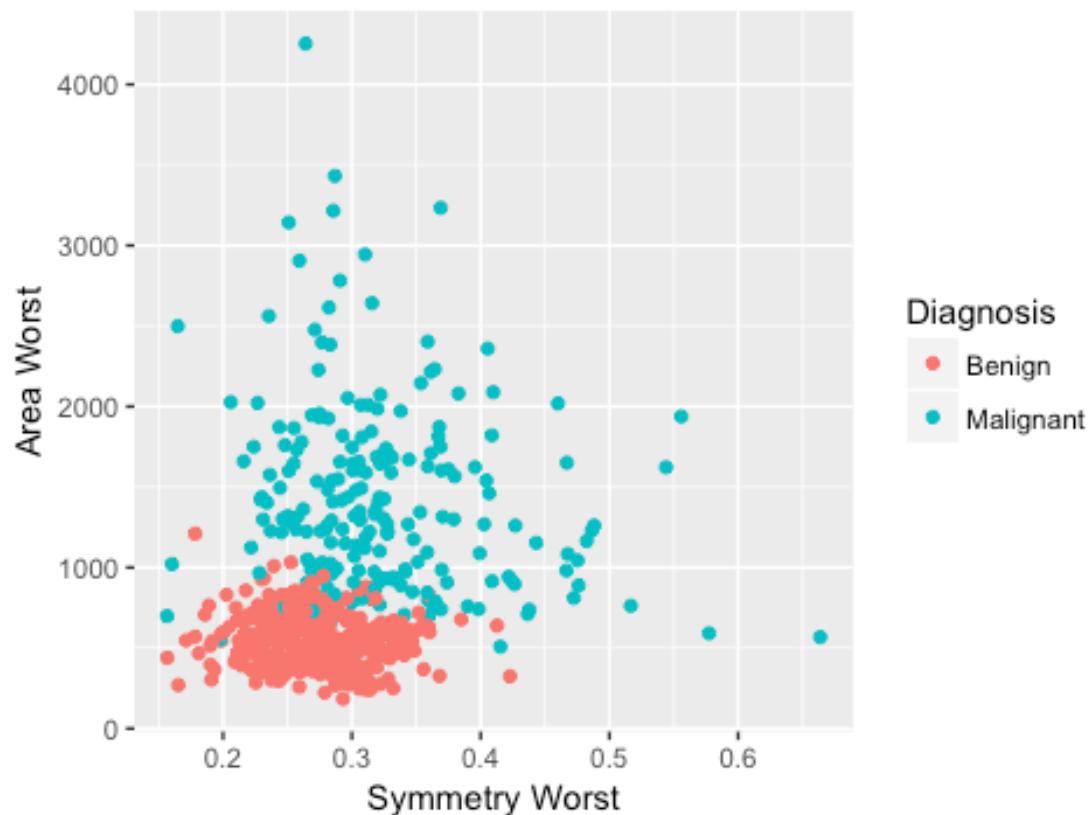
```
#perimeter_worst vs. symmetry_worst
symmetry_perimeter_worst_scatter <- ggplot(data = cancer, aes(x = symmetry_worst, y = perimeter_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Worst", y = "Perimeter Worst", title = "Perimeter Worst vs. Symmetry Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_perimeter_worst_scatter
```

Perimeter Worst vs. Symmetry Worst



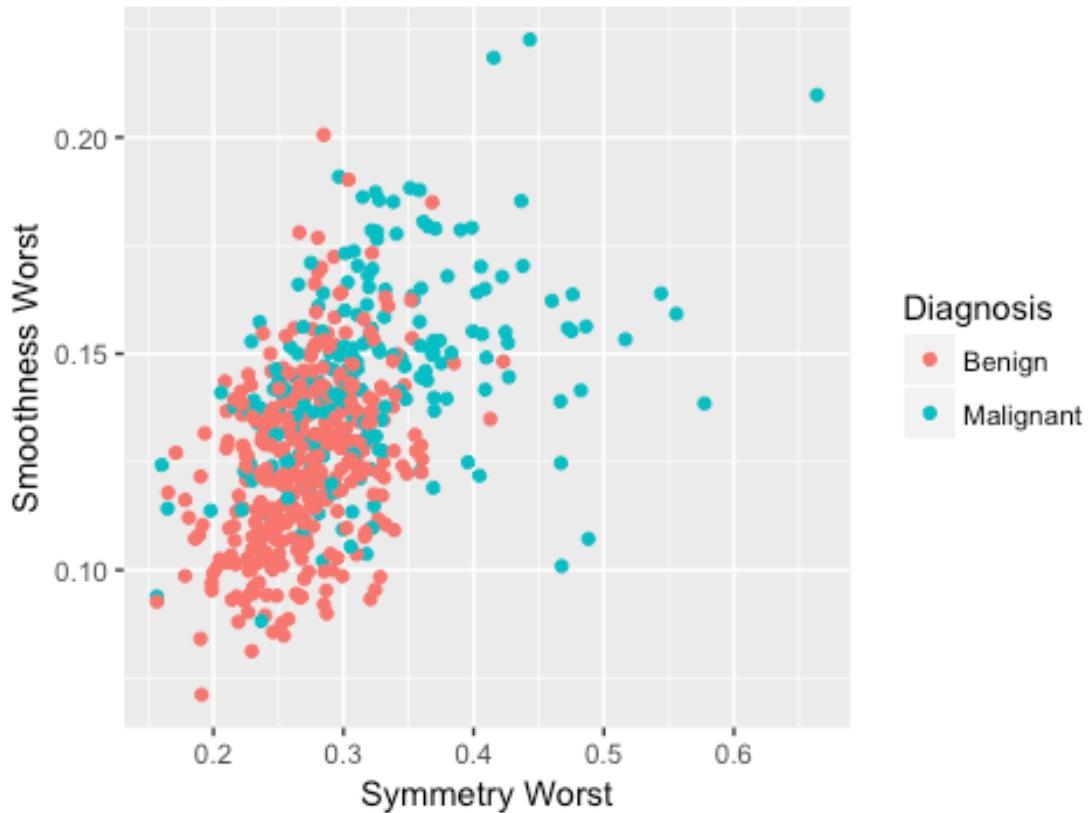
```
#area_worst vs. symmetry_worst
symmetry_area_worst_scatter <- ggplot(data = cancer, aes(x = symmetry_worst,
y = area_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Worst", y = "Area Worst", title = "Area Worst vs. Symmetry Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_area_worst_scatter
```

Area Worst vs. Symmetry Worst



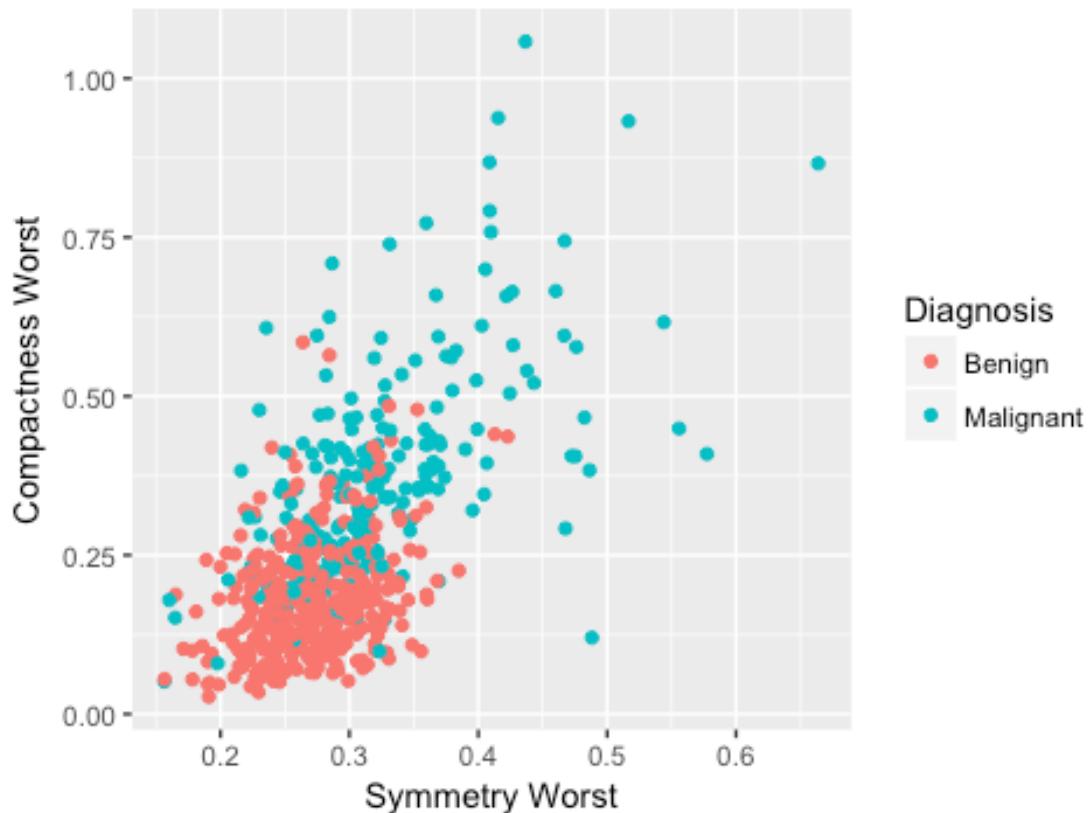
```
#smoothness_worst vs. symmetry_worst
symmetry_smoothness_worst_scatter <- ggplot(data = cancer, aes(x = symmetry_worst, y = smoothness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Worst", y = "Smoothness Worst", title = "Smoothness Worst vs. Symmetry Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_smoothness_worst_scatter
```

Smoothness Worst vs. Symmetry Worst



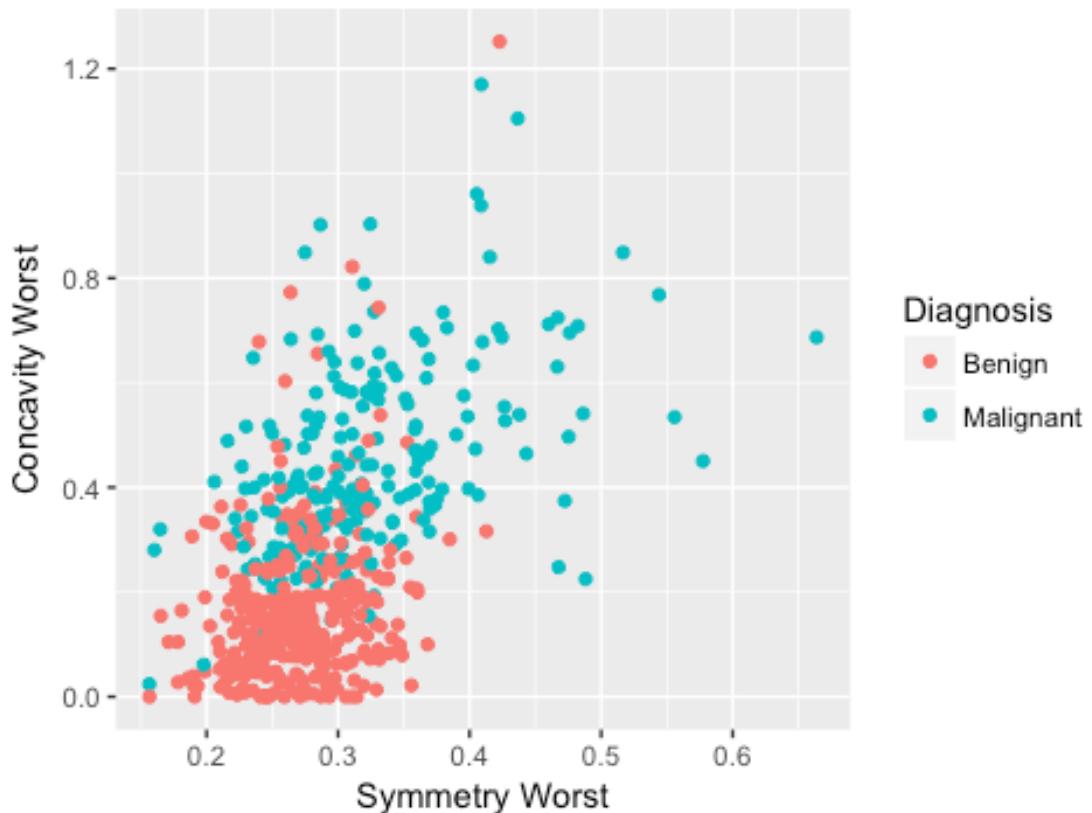
```
#compactness_worst vs. symmetry_worst
symmetry_compactness_worst_scatter <- ggplot(data = cancer, aes(x = symmetry_worst, y = compactness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Worst", y = "Compactness Worst", title =
  "Compactness Worst vs. Symmetry Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_compactness_worst_scatter
```

Compactness Worst vs. Symmetry Worst



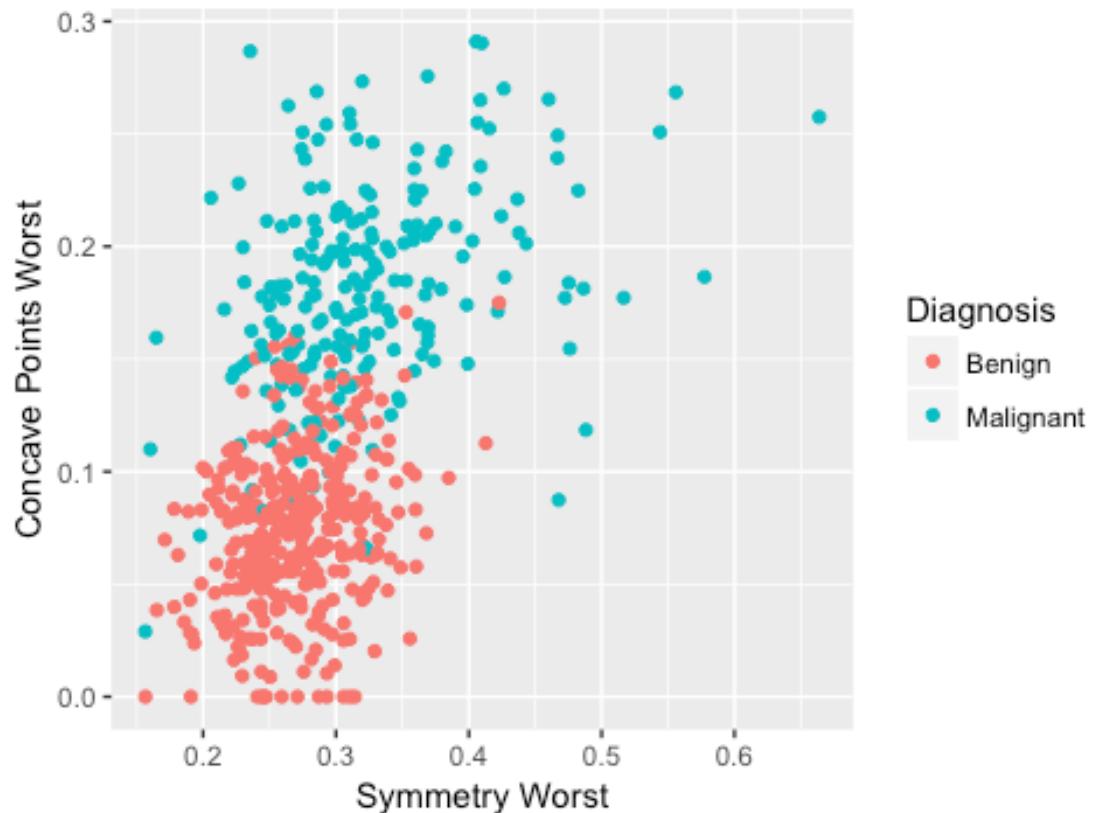
```
#concavity_worst vs. symmetry_worst
symmetry_concavity_worst_scatter <- ggplot(data = cancer, aes(x = symmetry_worst, y = concavity_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Worst", y = "Concavity Worst", title = "Concavity Worst vs. Symmetry Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_concavity_worst_scatter
```

Concavity Worst vs. Symmetry Worst



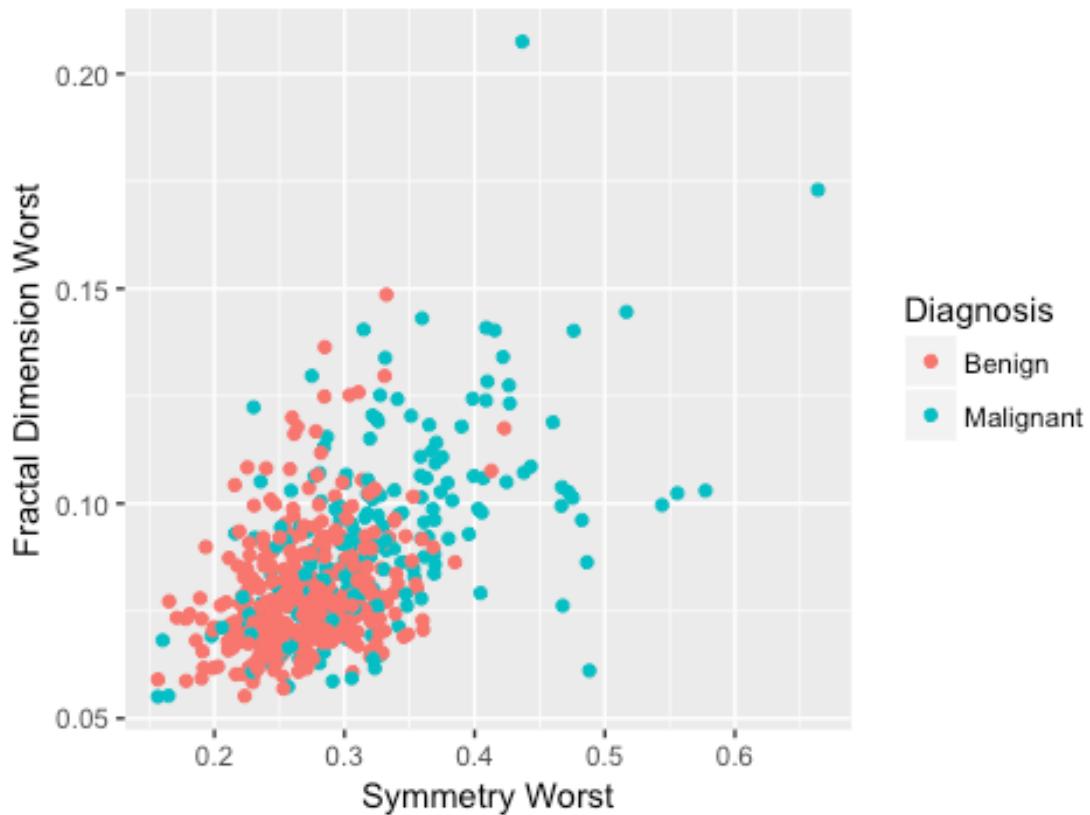
```
#concave.points_worst vs. symmetry_worst
symmetry_concave.points_worst_scatter <- ggplot(data = cancer, aes(x = symmetry_worst, y = concave.points_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Worst", y = "Concave Points Worst", title =
  "Concave Points Worst vs. Symmetry Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_concave.points_worst_scatter
```

Concave Points Worst vs. Symmetry Worst



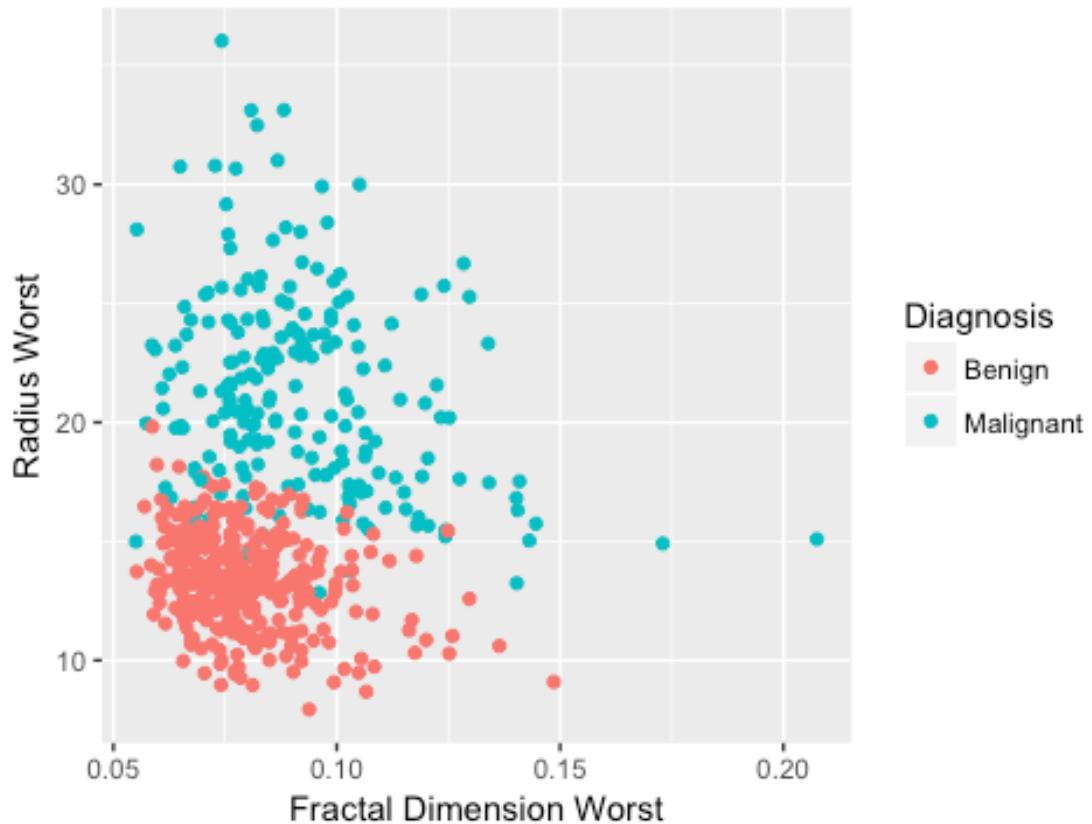
```
#fractal_dimension_worst vs. symmetry_worst
symmetry_fractal_dimension_worst_scatter <- ggplot(data = cancer, aes(x = symmetry_worst, y = fractal_dimension_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Worst", y = "Fractal Dimension Worst", title = "Fractal Dimension Worst vs. Symmetry Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_fractal_dimension_worst_scatter
```

Fractal Dimension Worst vs. Symmetry Worst



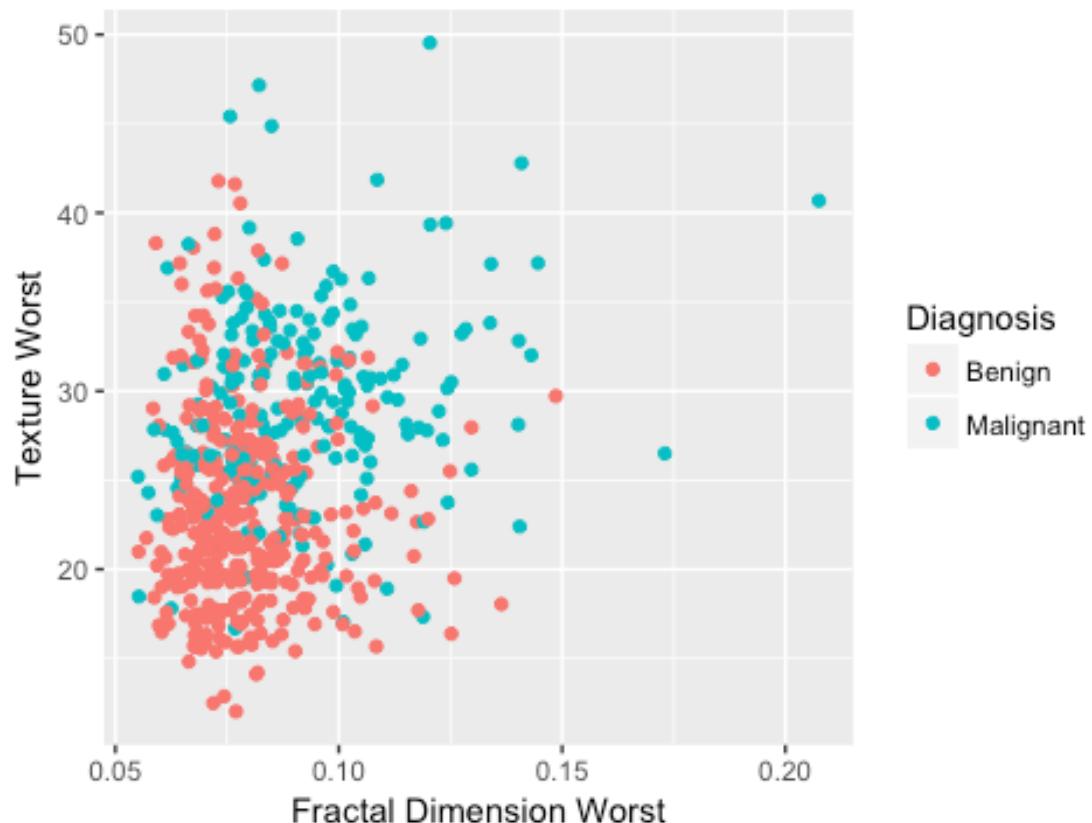
```
#radius_worst vs. fractal_dimension_worst
fractal_dimension_radius_worst_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_worst, y = radius_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Worst", y = "Radius Worst", title = "Radius Worst vs. Fractal Dimension Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_radius_worst_scatter
```

Radius Worst vs. Fractal Dimension Worst



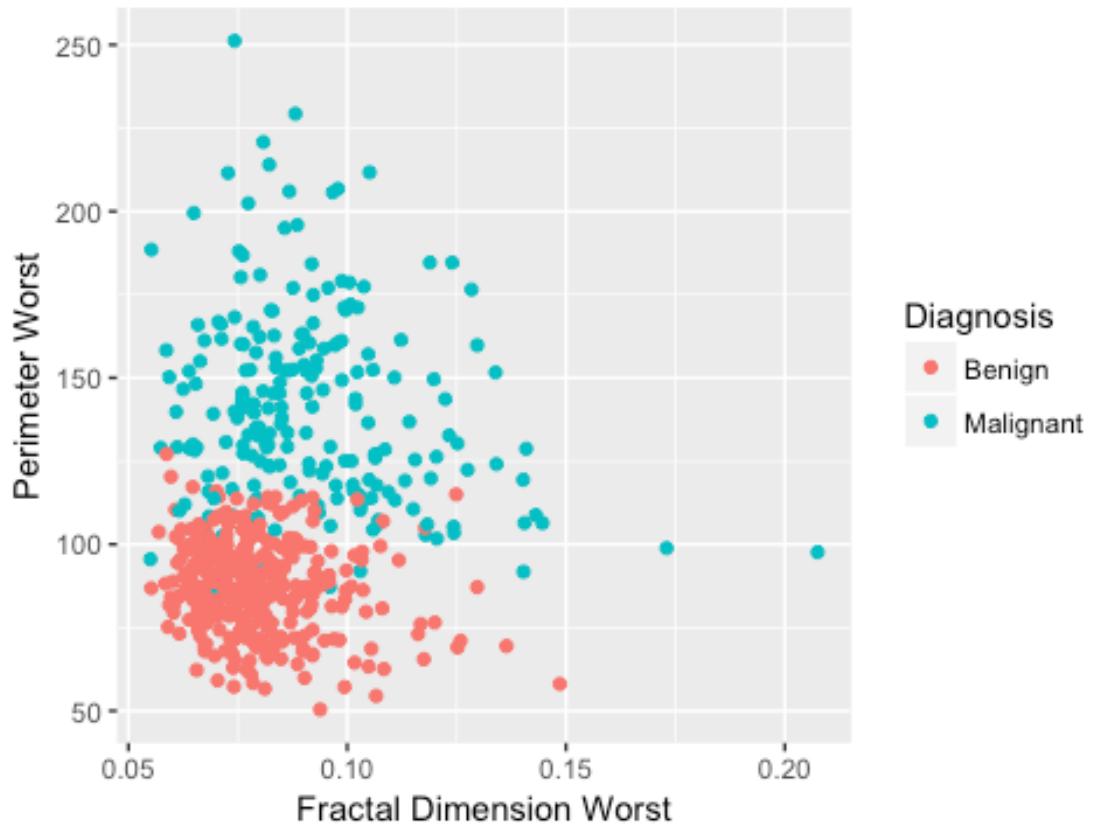
```
#texture_worst vs. fractal_dimension_worst
fractal_dimension_texture_worst_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_worst, y = texture_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Worst", y = "Texture Worst", title = "Texture Worst vs. Fractal Dimension Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_texture_worst_scatter
```

Texture Worst vs. Fractal Dimension Worst



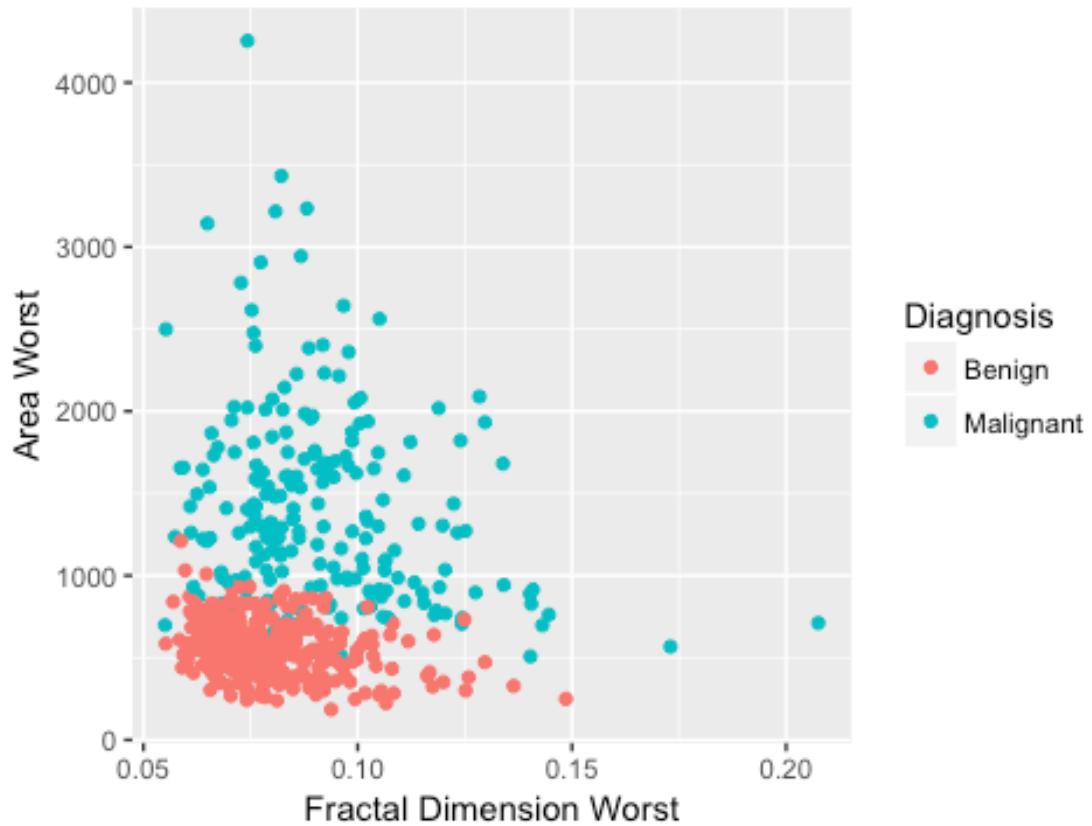
```
#perimeter_worst vs. fractal_dimension_worst
fractal_dimension_perimeter_worst_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_worst, y = perimeter_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Worst", y = "Perimeter Worst", title = "Perimeter Worst vs. Fractal Dimension Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_perimeter_worst_scatter
```

Perimeter Worst vs. Fractal Dimension Worst



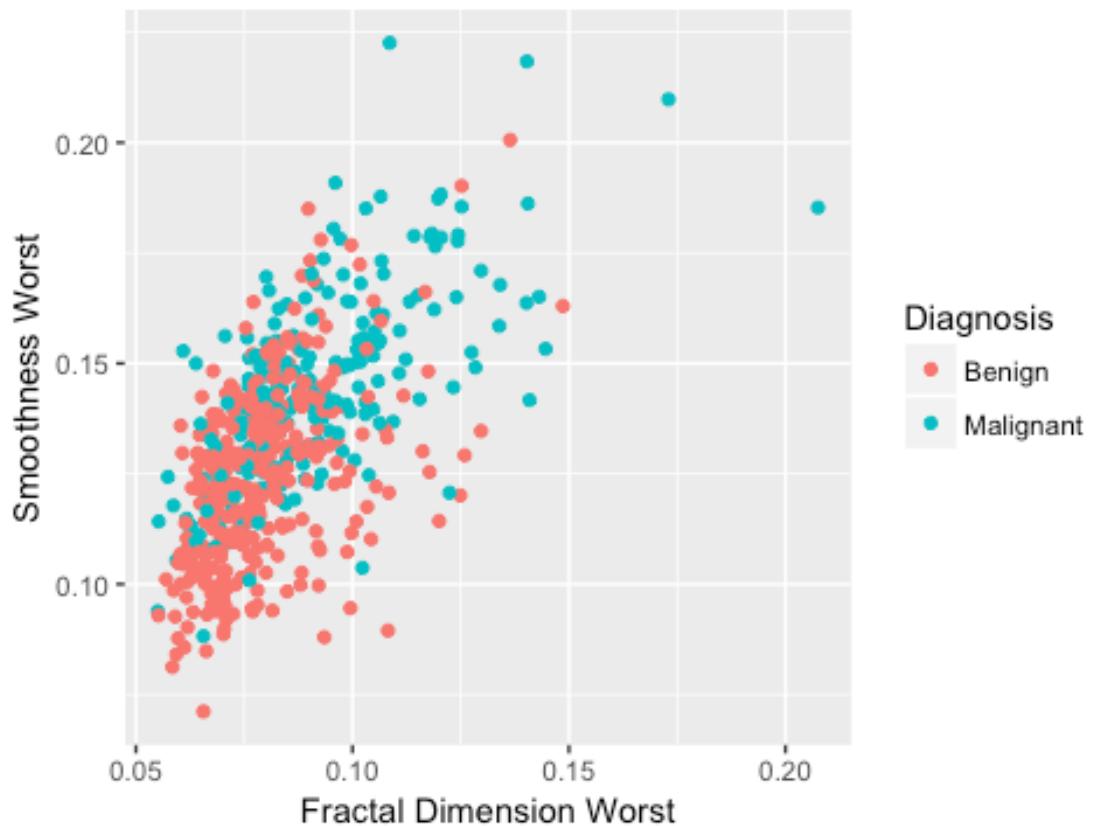
```
#area_worst vs. fractal_dimension_worst
fractal_dimension_area_worst_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_worst, y = area_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Worst", y = "Area Worst", title =
  "Area Worst vs. Fractal Dimension Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_area_worst_scatter
```

Area Worst vs. Fractal Dimension Worst



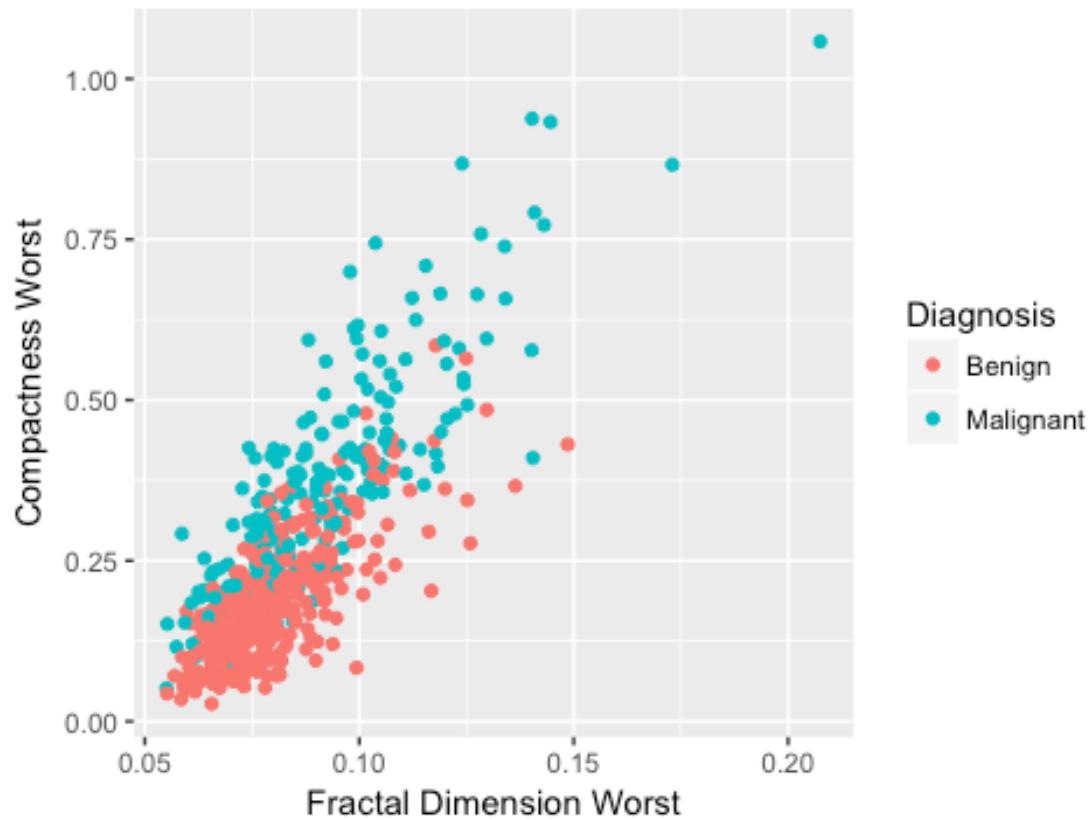
```
#smoothness_worst vs. fractal_dimension_worst
fractal_dimension_smoothness_worst_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_worst, y = smoothness_worst, color = as.factor(diagnosis)))
+
  geom_point() + labs(x = "Fractal Dimension Worst", y = "Smoothness Worst",
  title = "Smoothness Worst vs. Fractal Dimension Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_smoothness_worst_scatter
```

Smoothness Worst vs. Fractal Dimension Worst



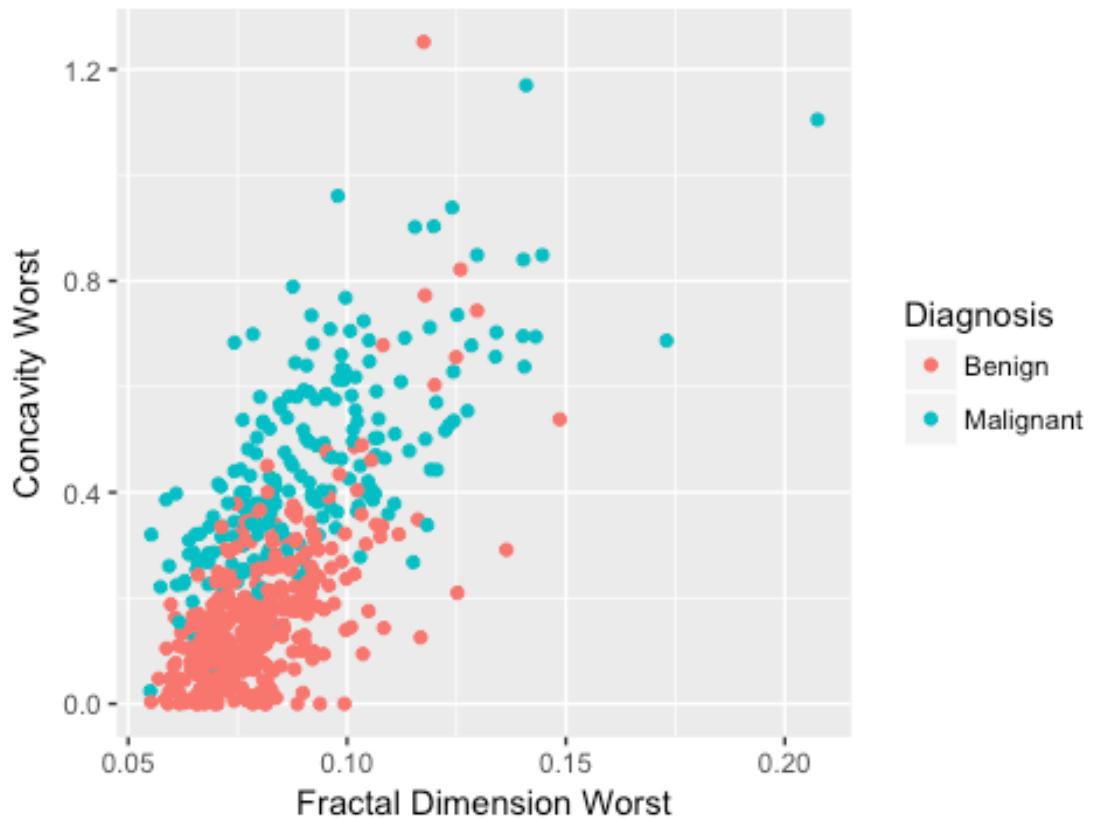
```
#compactness_worst vs. fractal_dimension_worst
fractal_dimension_compactness_worst_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_worst, y = compactness_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Worst", y = "Compactness Worst",
  title = "Compactness Worst vs. Fractal Dimension Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_compactness_worst_scatter
```

Compactness Worst vs. Fractal Dimension Worst



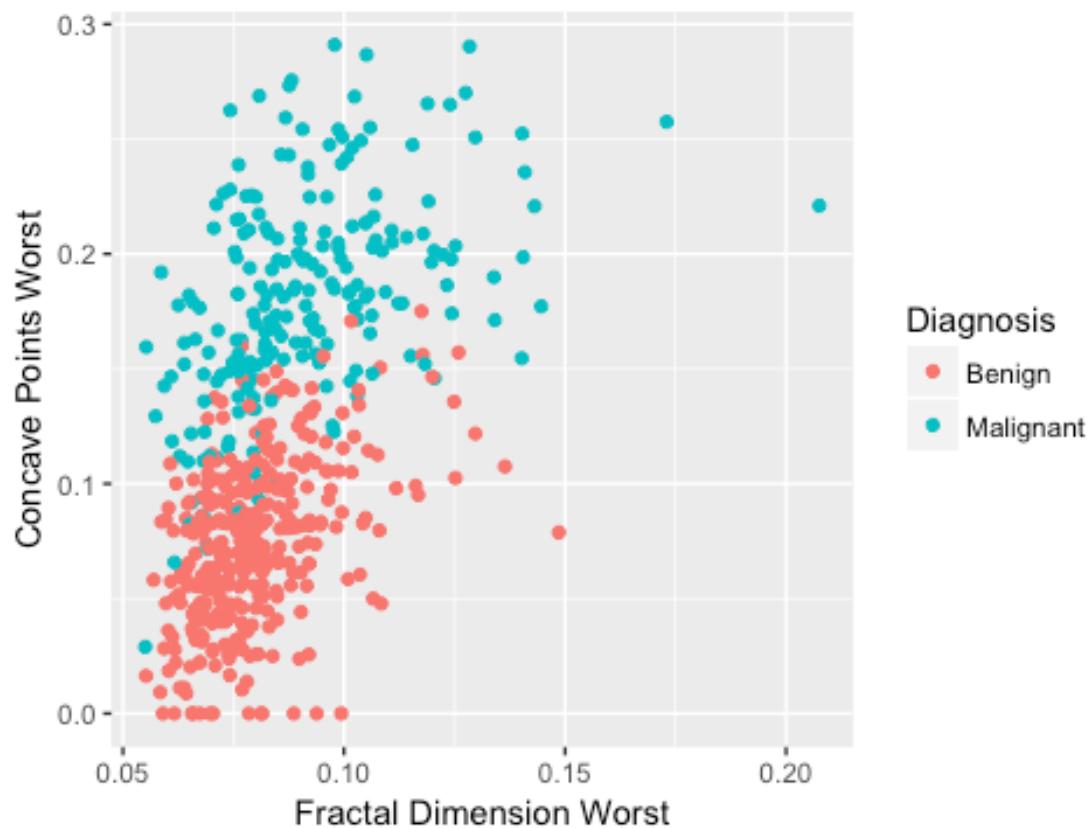
```
#concavity_worst vs. fractal_dimension_worst
fractal_dimension_concavity_worst_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_worst, y = concavity_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Worst", y = "Concavity Worst", title = "Concavity Worst vs. Fractal Dimension Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_concavity_worst_scatter
```

Concavity Worst vs. Fractal Dimension Worst



```
#concave.points_worst vs. fractal_dimension_worst
fractal_dimension_concave.points_worst_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_worst, y = concave.points_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Worst", y = "Concave Points Worst", title = "Concave Points Worst vs. Fractal Dimension Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_concave.points_worst_scatter
```

Concave Points Worst vs. Fractal Dimension Worst



```
#symmetry_worst vs. fractal_dimension_worst
fractal_dimension_symmetry_worst_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_worst, y = symmetry_worst, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Worst", y = "Symmetry Worst", title = "Symmetry Worst vs. Fractal Dimension Worst") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_symmetry_worst_scatter
```

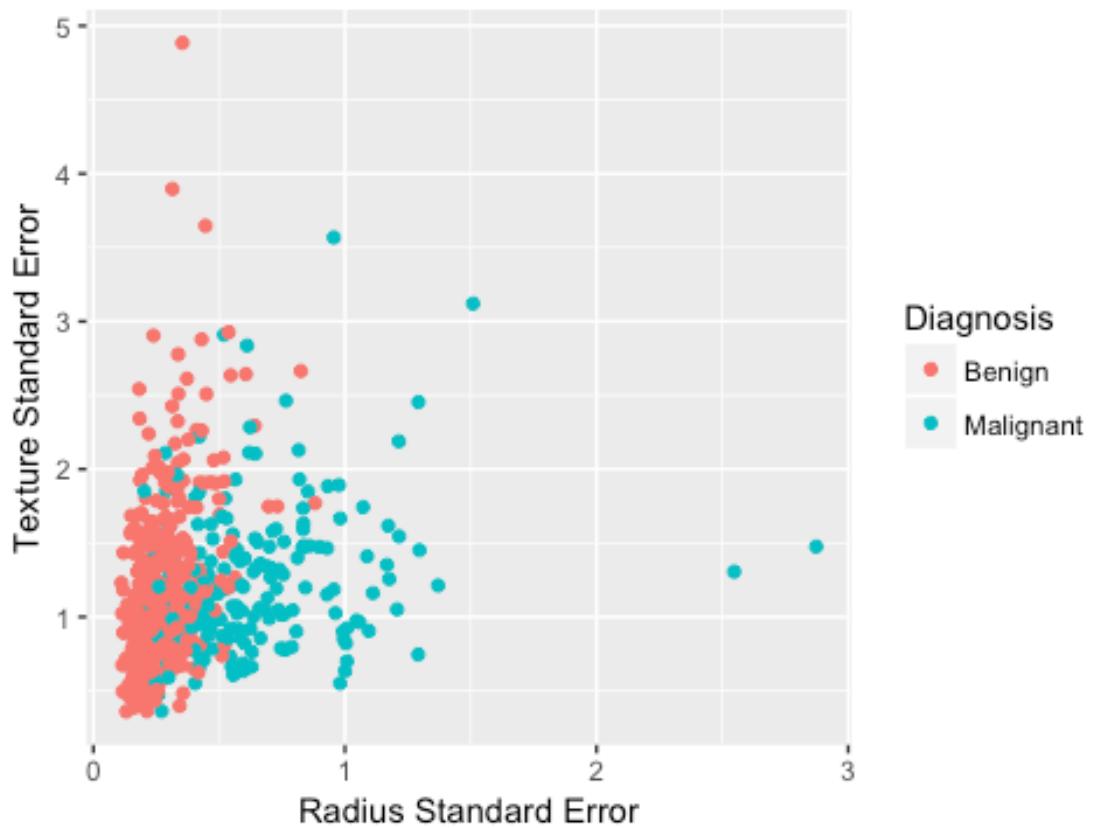
Symmetry Worst vs. Fractal Dimension Worst



```
#Standard Error variable relationships
```

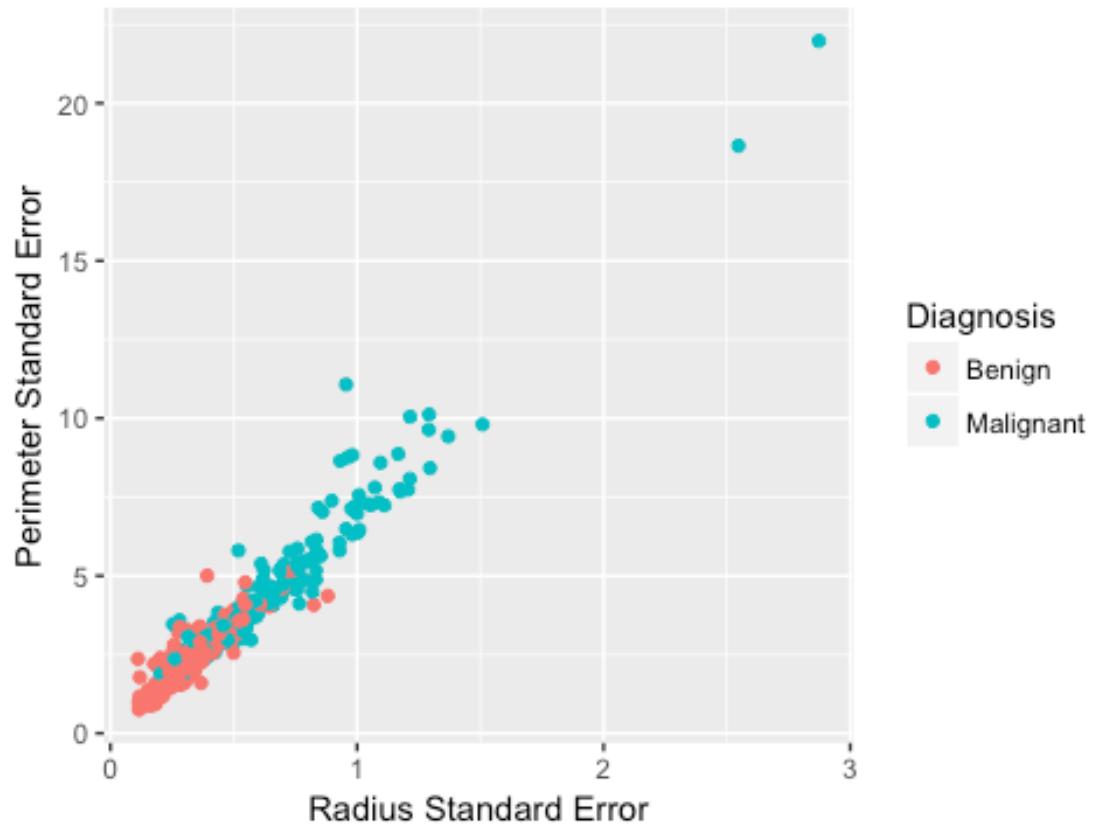
```
#texture_se vs. radius_se
radius_texture_se_scatter <- ggplot(data = cancer, aes(x = radius_se, y = texture_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Standard Error", y = "Texture Standard Error", title = "Texture Standard Error vs. Radius Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_texture_se_scatter
```

Texture Standard Error vs. Radius Standard Error



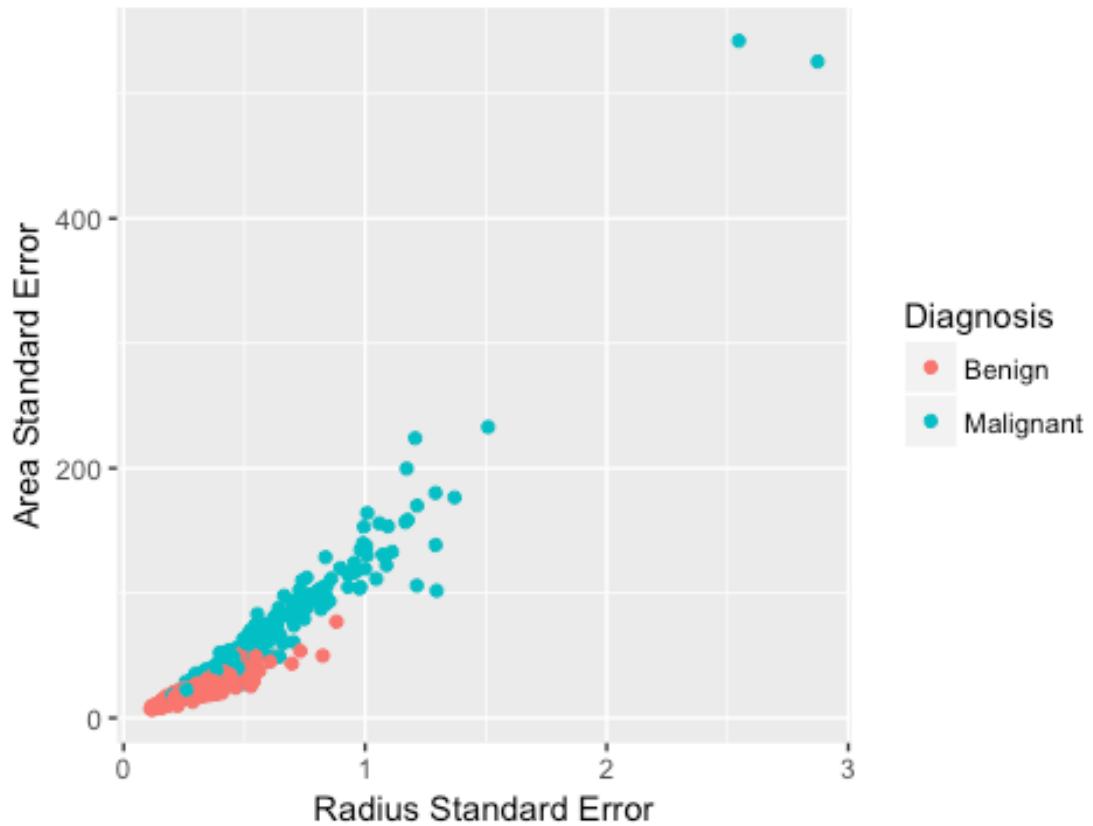
```
#perimeter_se vs. radius_se
radius_perimeter_se_scatter <- ggplot(data = cancer, aes(x = radius_se, y = perimeter_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Standard Error", y = "Perimeter Standard Error", title = "Perimeter Standard Error vs. Radius Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_perimeter_se_scatter
```

Perimeter Standard Error vs. Radius Standard Error



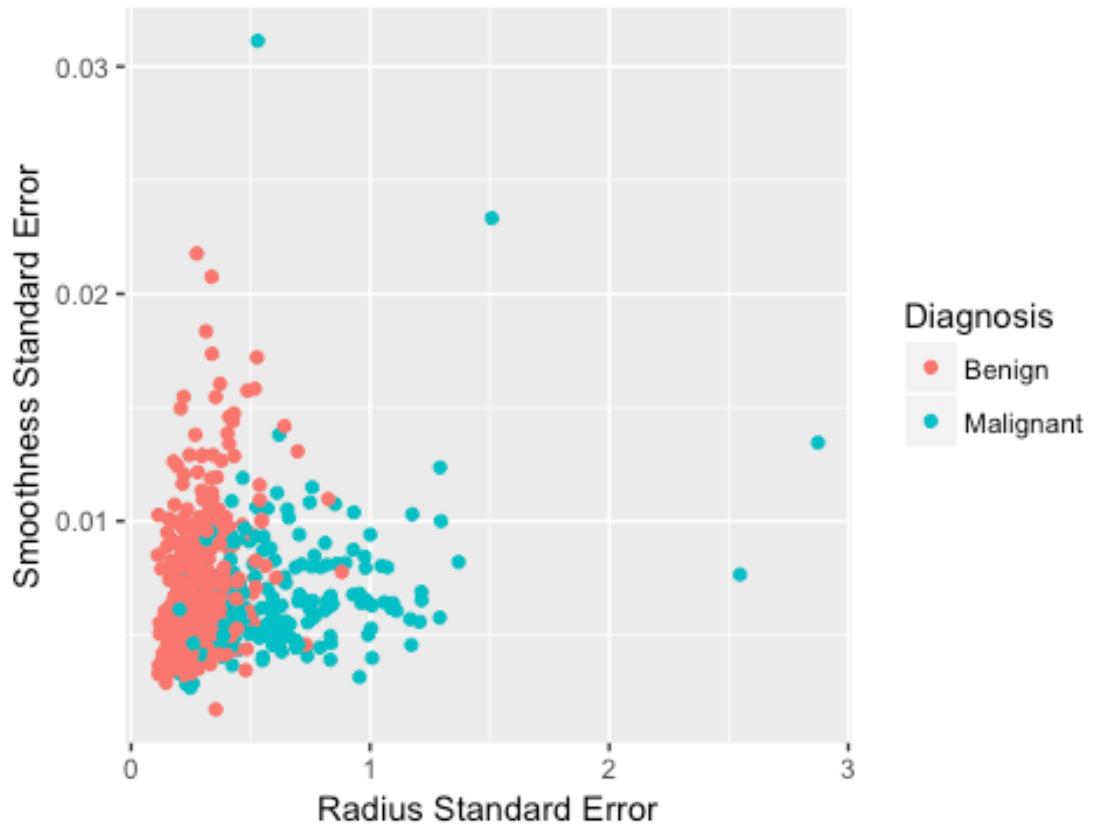
```
#area_se vs. radius_se
radius_area_se_scatter <- ggplot(data = cancer, aes(x = radius_se, y = area_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Standard Error", y = "Area Standard Error",
  title = "Area Standard Error vs. Radius Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_area_se_scatter
```

Area Standard Error vs. Radius Standard Error



```
#smoothness_se vs. radius_se
radius_smoothness_se_scatter <- ggplot(data = cancer, aes(x = radius_se, y = smoothness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Standard Error", y = "Smoothness Standard Error", title = "Smoothness Standard Error vs. Radius Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_smoothness_se_scatter
```

Smoothness Standard Error vs. Radius Standard Error



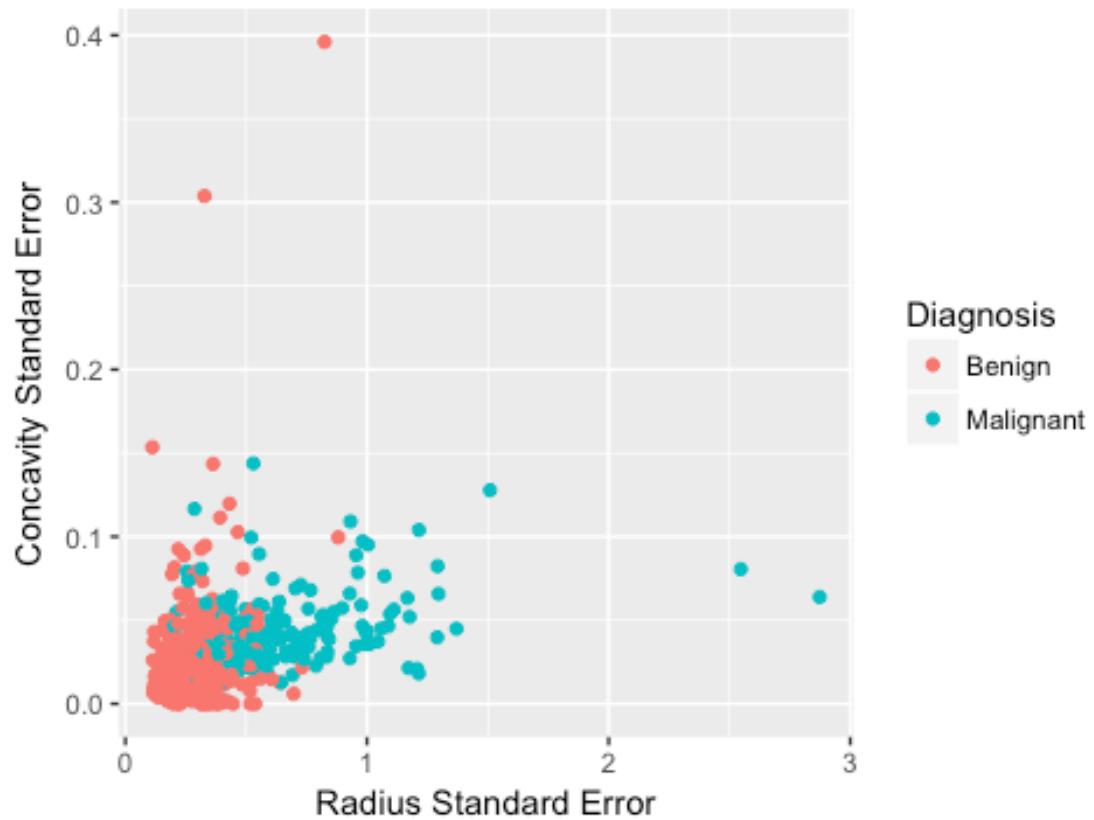
```
#compactness_se vs. radius_se
radius_compactness_se_scatter <- ggplot(data = cancer, aes(x = radius_se, y = compactness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Standard Error", y = "Compactness Standard Error", title = "Compactness Standard Error vs. Radius Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_compactness_se_scatter
```

Compactness Standard Error vs. Radius Standard Error



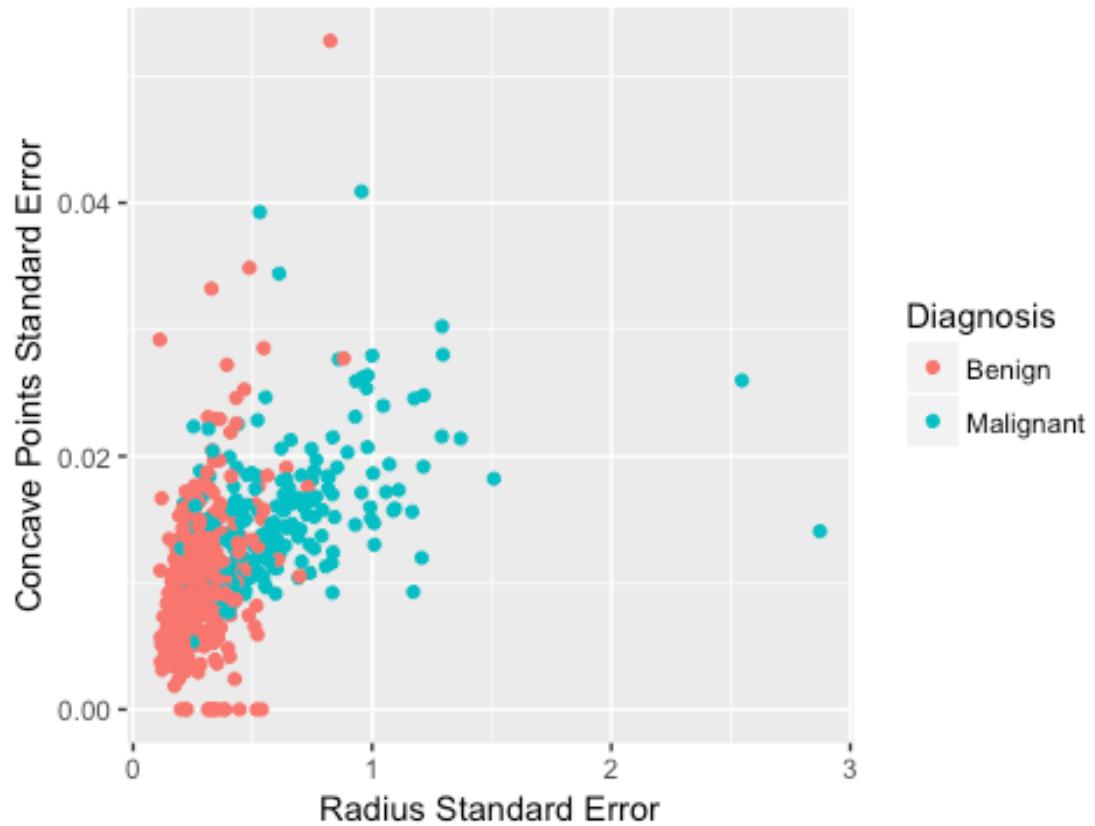
```
#concavity_se vs. radius_se
radius_concavity_se_scatter <- ggplot(data = cancer, aes(x = radius_se, y = concavity_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Standard Error", y = "Concavity Standard Error", title = "Concavity Standard Error vs. Radius Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_concavity_se_scatter
```

Concavity Standard Error vs. Radius Standard Error



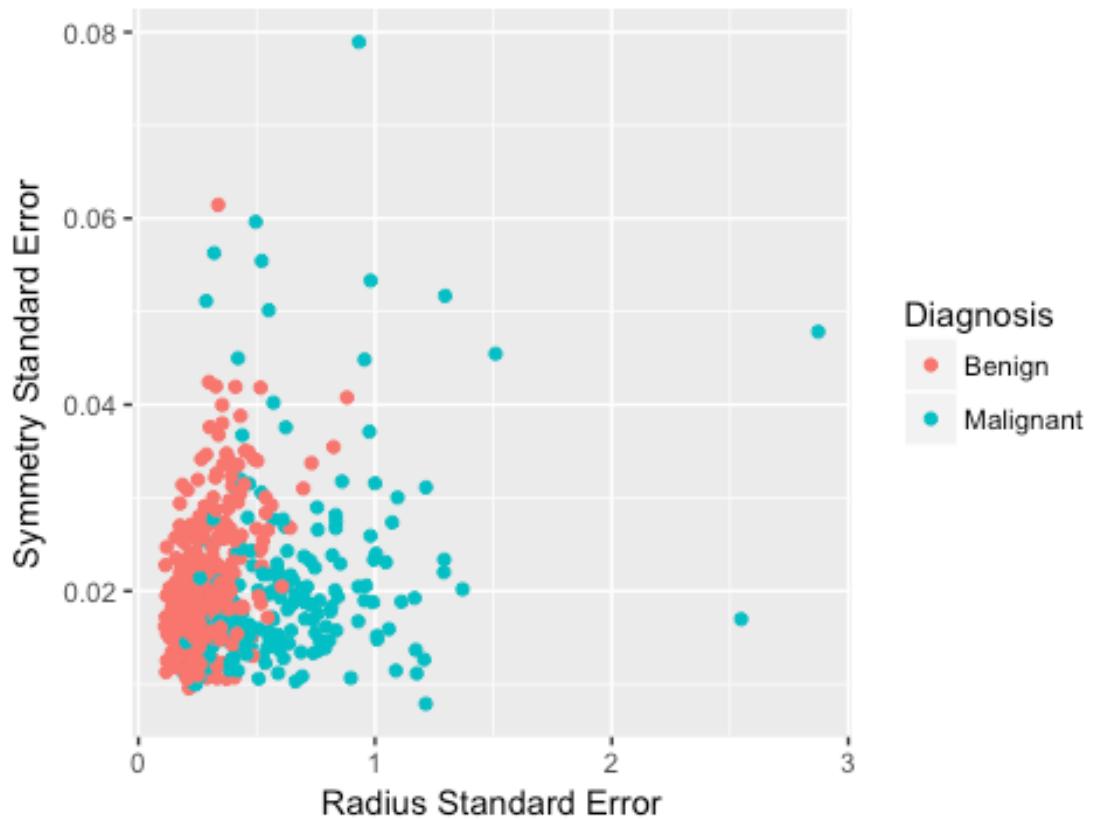
```
#concave.points_se vs. radius_se
radius_concave.points_se_scatter <- ggplot(data = cancer, aes(x = radius_se,
y = concave.points_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Standard Error", y = "Concave Points Standard Error", title = "Concave Points Standard Error vs. Radius Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_concave.points_se_scatter
```

Concave Points Standard Error vs. Radius Standard Error



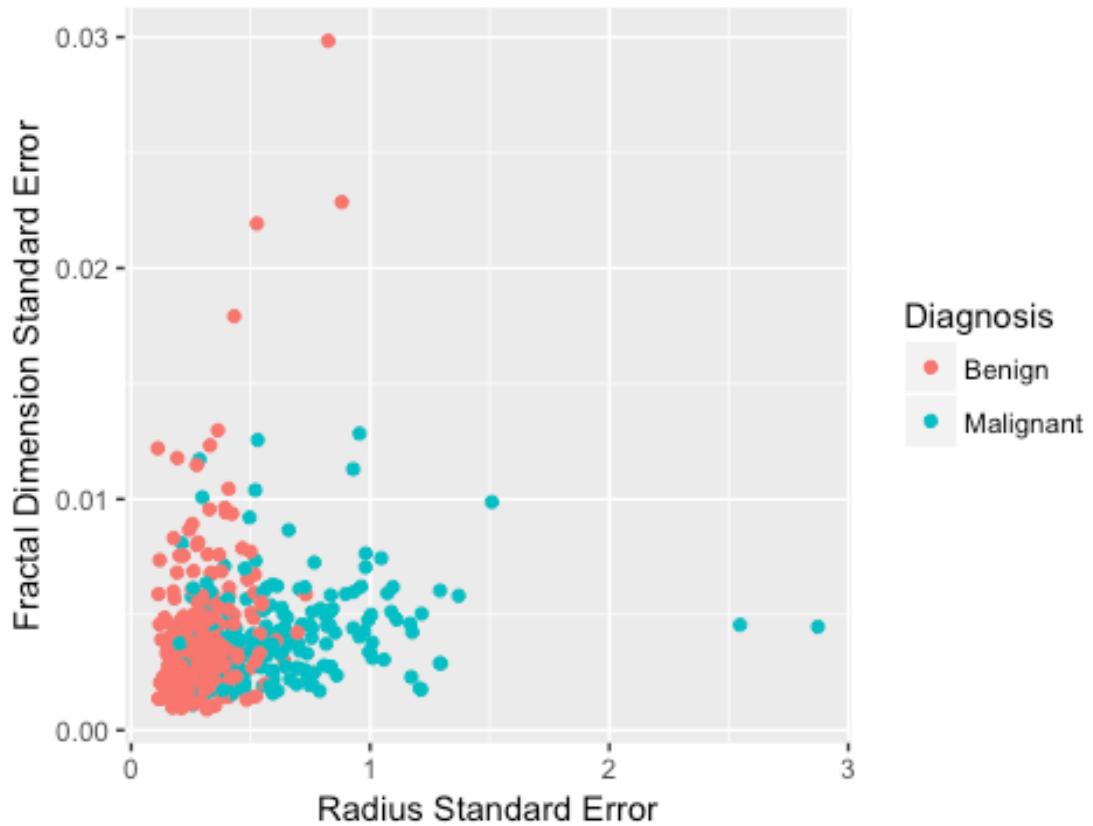
```
#symmetry_se vs. radius_se
radius_symmetry_se_scatter <- ggplot(data = cancer, aes(x = radius_se, y = symmetry_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Standard Error", y = "Symmetry Standard Error", title = "Symmetry Standard Error vs. Radius Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_symmetry_se_scatter
```

Symmetry Standard Error vs. Radius Standard Error



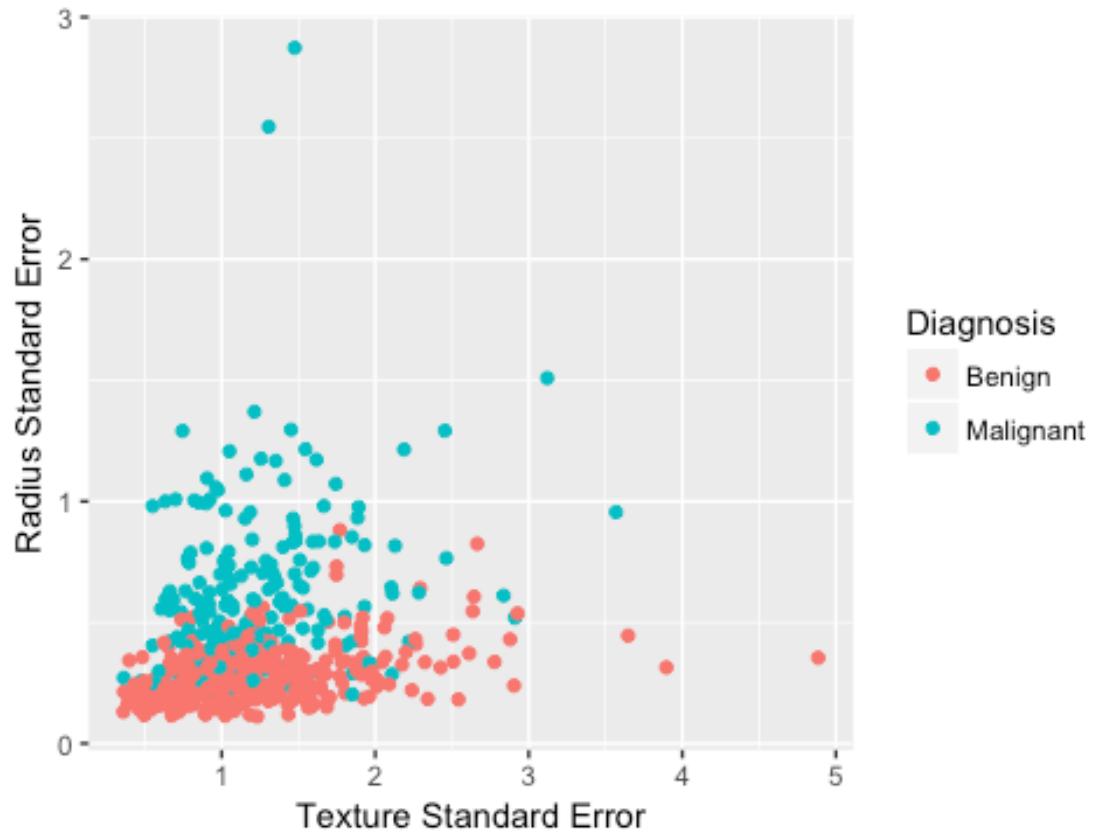
```
#fractal_dimension_se vs. radius_se
radius_fractal_dimension_se_scatter <- ggplot(data = cancer, aes(x = radius_se, y = fractal_dimension_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Radius Standard Error", y = "Fractal Dimension Standard Error", title = "Fractal Dimension Standard Error vs. Radius Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
radius_fractal_dimension_se_scatter
```

Fractal Dimension Standard Error vs. Radius Standard



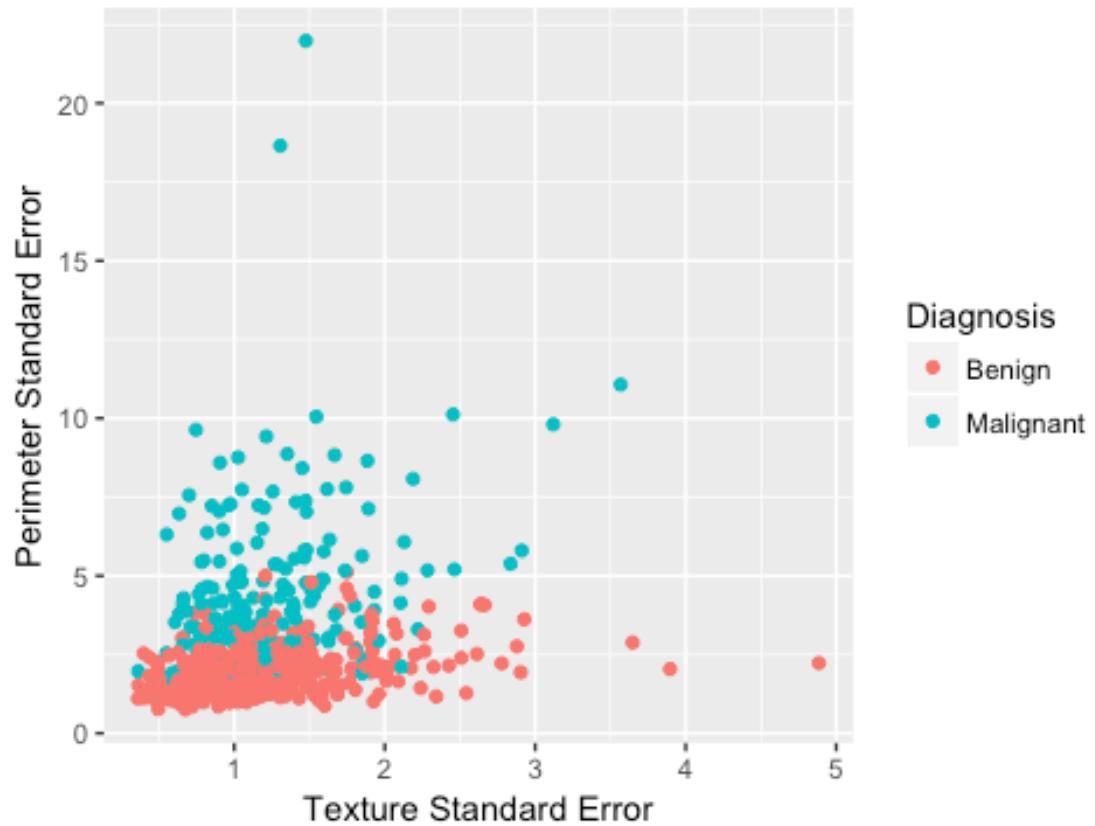
```
#radius_se vs. texture_se
texture_radius_se_scatter <- ggplot(data = cancer, aes(x = texture_se, y = radius_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Standard Error", y = "Radius Standard Error", title = "Radius Standard Error vs. Texture Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_radius_se_scatter
```

Radius Standard Error vs. Texture Standard Error



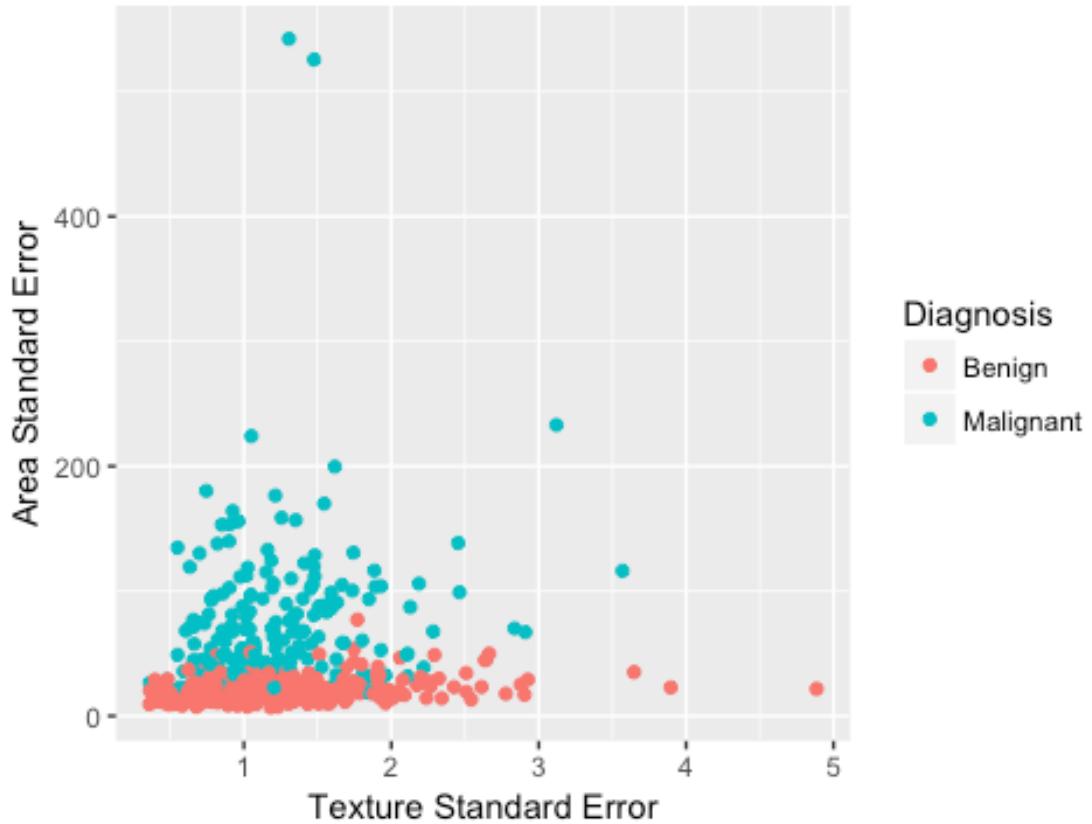
```
#perimeter_se vs. texture_se
texture_perimeter_se_scatter <- ggplot(data = cancer, aes(x = texture_se, y = perimeter_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Standard Error", y = "Perimeter Standard Error", title = "Perimeter Standard Error vs. Texture Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_perimeter_se_scatter
```

Perimeter Standard Error vs. Texture Standard Error



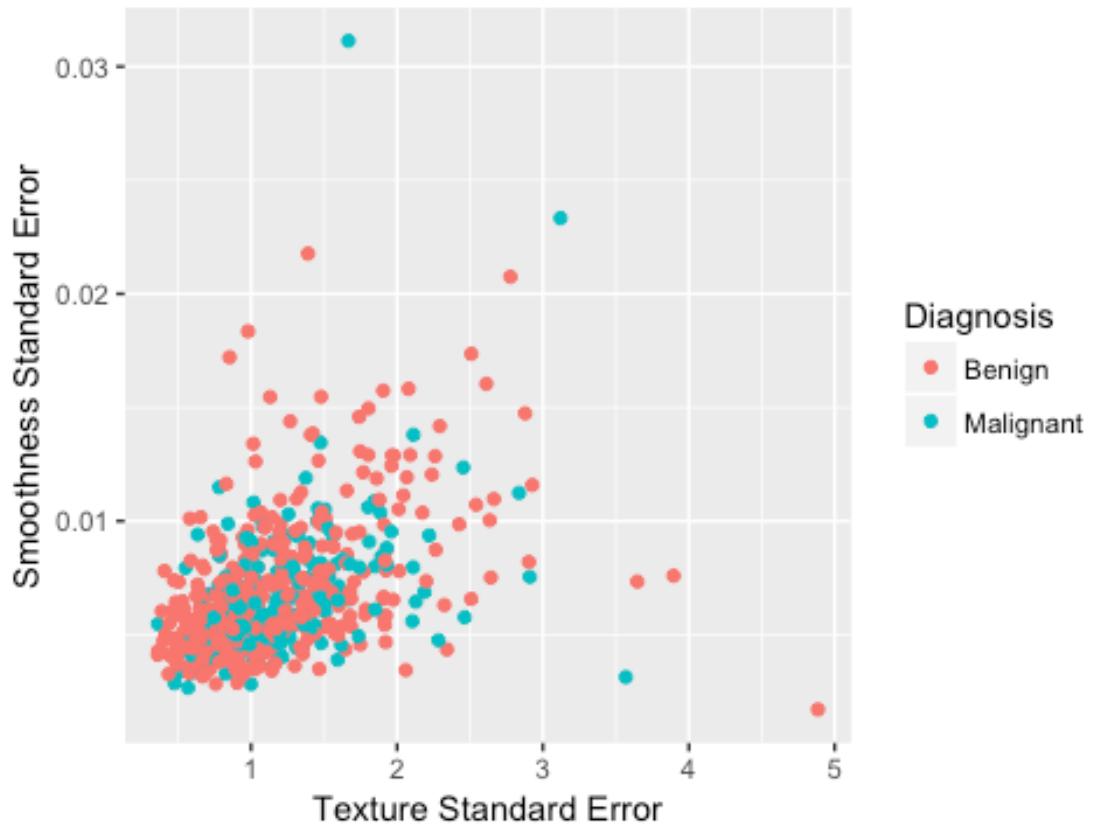
```
#area_se vs. texture_se
texture_area_se_scatter <- ggplot(data = cancer, aes(x = texture_se, y = area_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Standard Error", y = "Area Standard Error",
  title = "Area Standard Error vs. Texture Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_area_se_scatter
```

Area Standard Error vs. Texture Standard Error



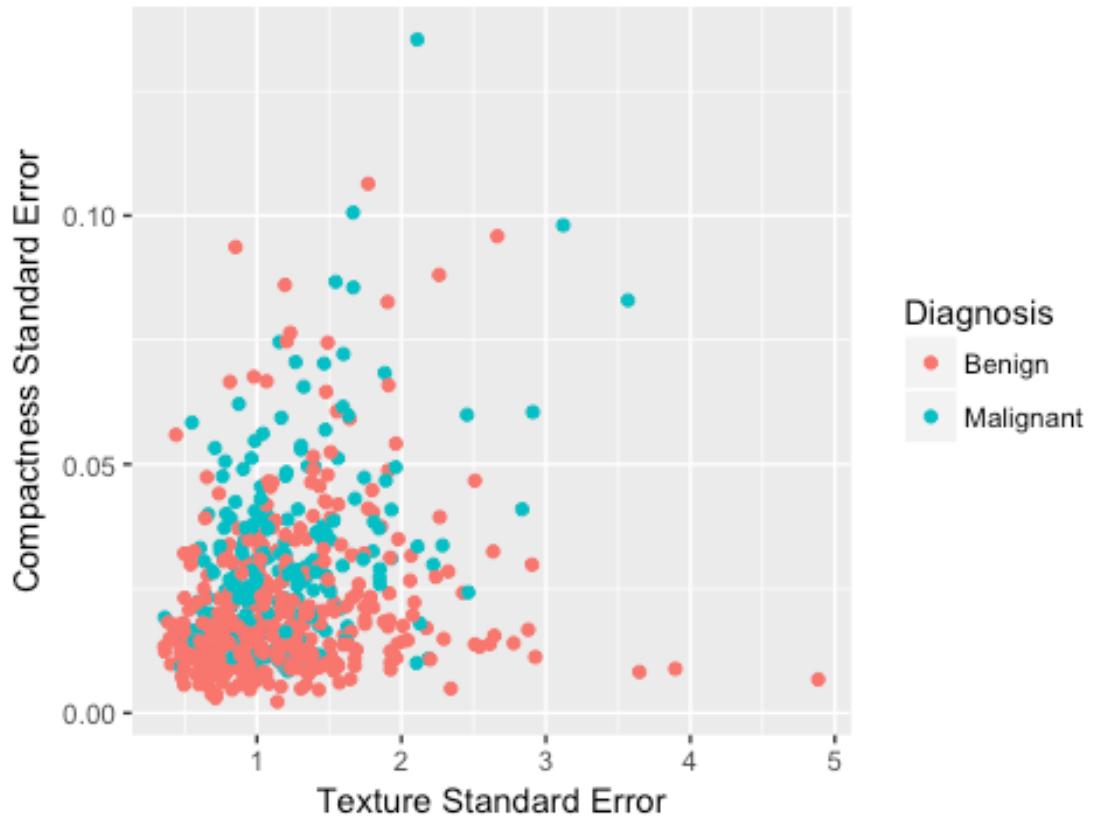
```
#smoothness_se vs. texture_se
texture_smoothness_se_scatter <- ggplot(data = cancer, aes(x = texture_se, y = smoothness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Standard Error", y = "Smoothness Standard Error", title = "Smoothness Standard Error vs. Texture Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_smoothness_se_scatter
```

Smoothness Standard Error vs. Texture Standard Error



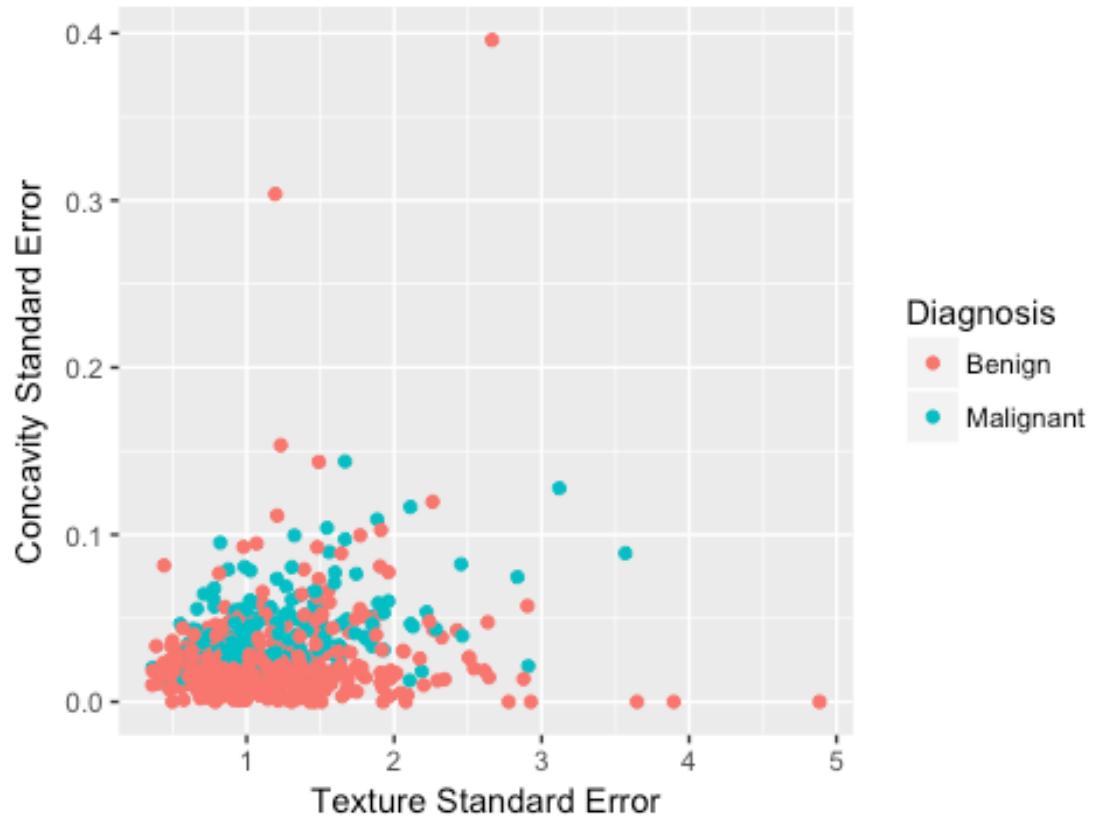
```
#compactness_se vs. texture_se
texture_compactness_se_scatter <- ggplot(data = cancer, aes(x = texture_se, y = compactness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Standard Error", y = "Compactness Standard Error", title = "Compactness Standard Error vs. Texture Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_compactness_se_scatter
```

Compactness Standard Error vs. Texture Standard Err



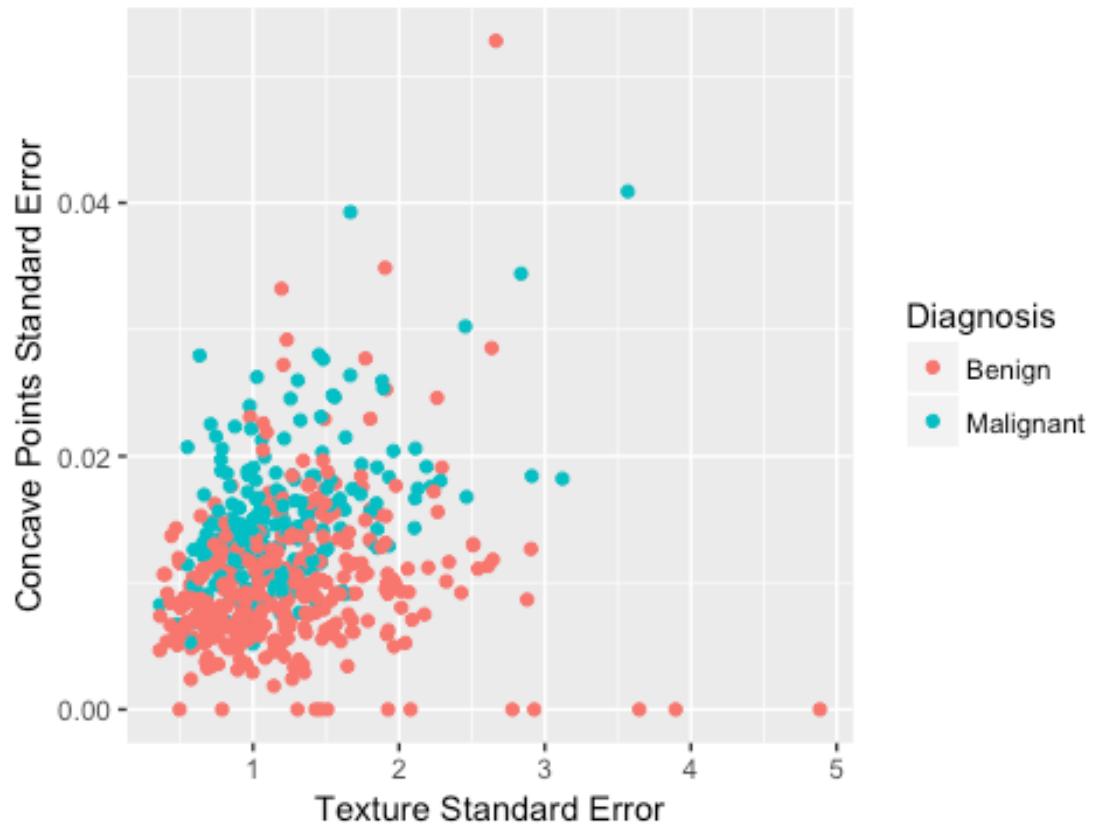
```
#concavity_se vs. texture_se
texture_concavity_se_scatter <- ggplot(data = cancer, aes(x = texture_se, y = concavity_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Standard Error", y = "Concavity Standard Error", title = "Concavity Standard Error vs. Texture Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_concavity_se_scatter
```

Concavity Standard Error vs. Texture Standard Error



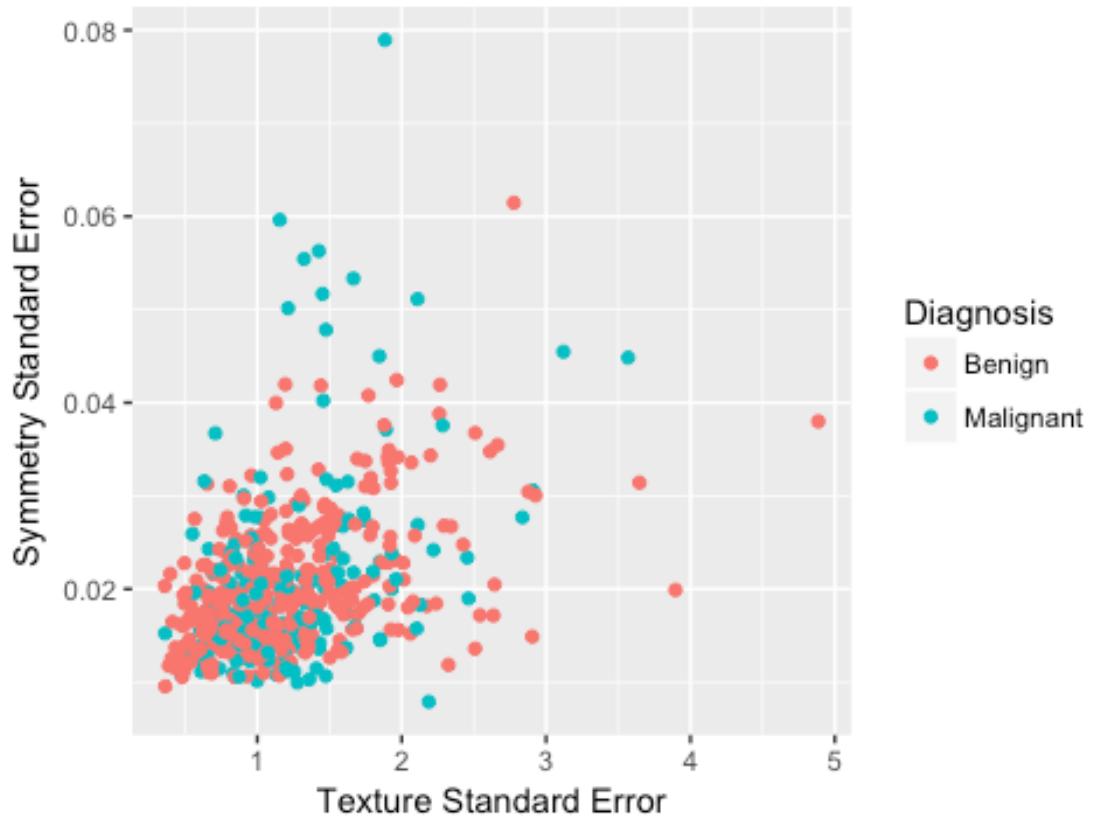
```
#concave.points_se vs. texture_se
texture_concave.points_se_scatter <- ggplot(data = cancer, aes(x = texture_se,
, y = concave.points_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Standard Error", y = "Concave Points Standard Error", title = "Concave Points Standard Error vs. Texture Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_concave.points_se_scatter
```

Concave Points Standard Error vs. Texture Standard E

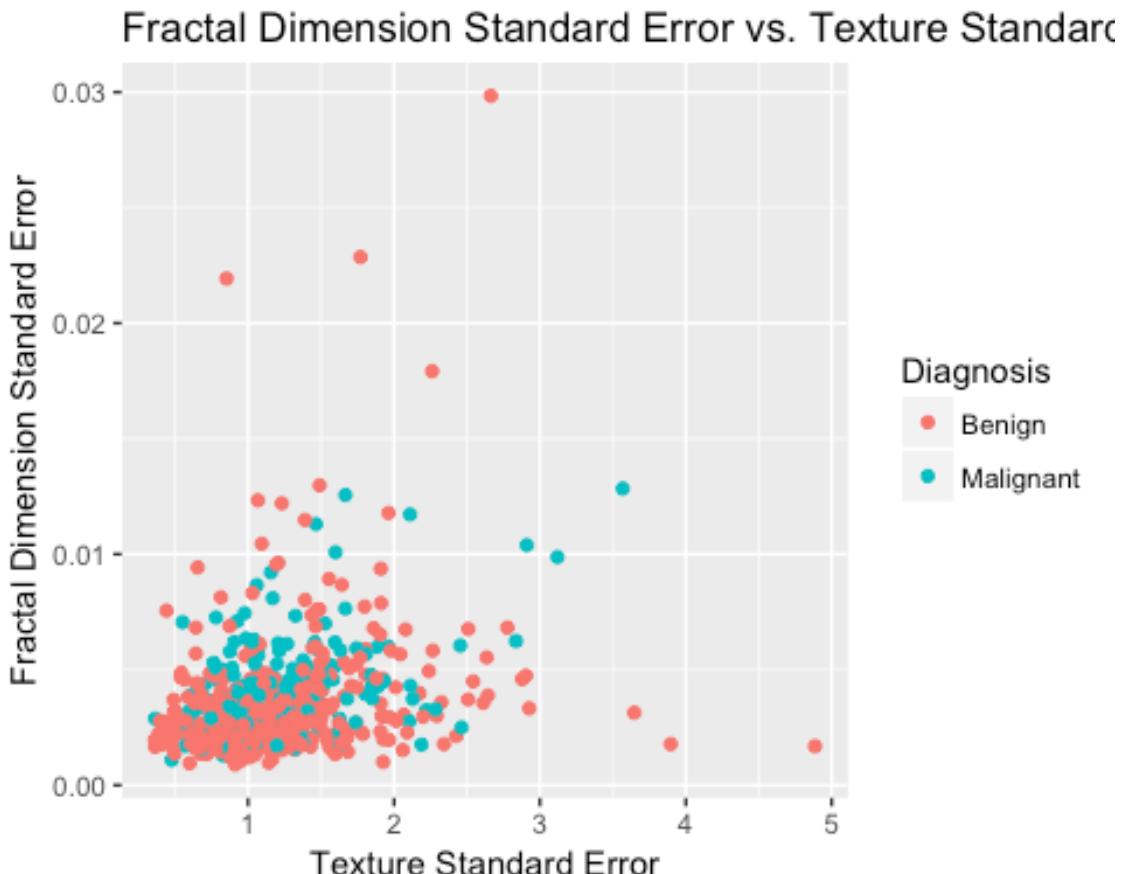


```
#symmetry_se vs. texture_se
texture_symmetry_se_scatter <- ggplot(data = cancer, aes(x = texture_se, y = symmetry_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Standard Error", y = "Symmetry Standard Error", title = "Symmetry Standard Error vs. Texture Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_symmetry_se_scatter
```

Symmetry Standard Error vs. Texture Standard Error

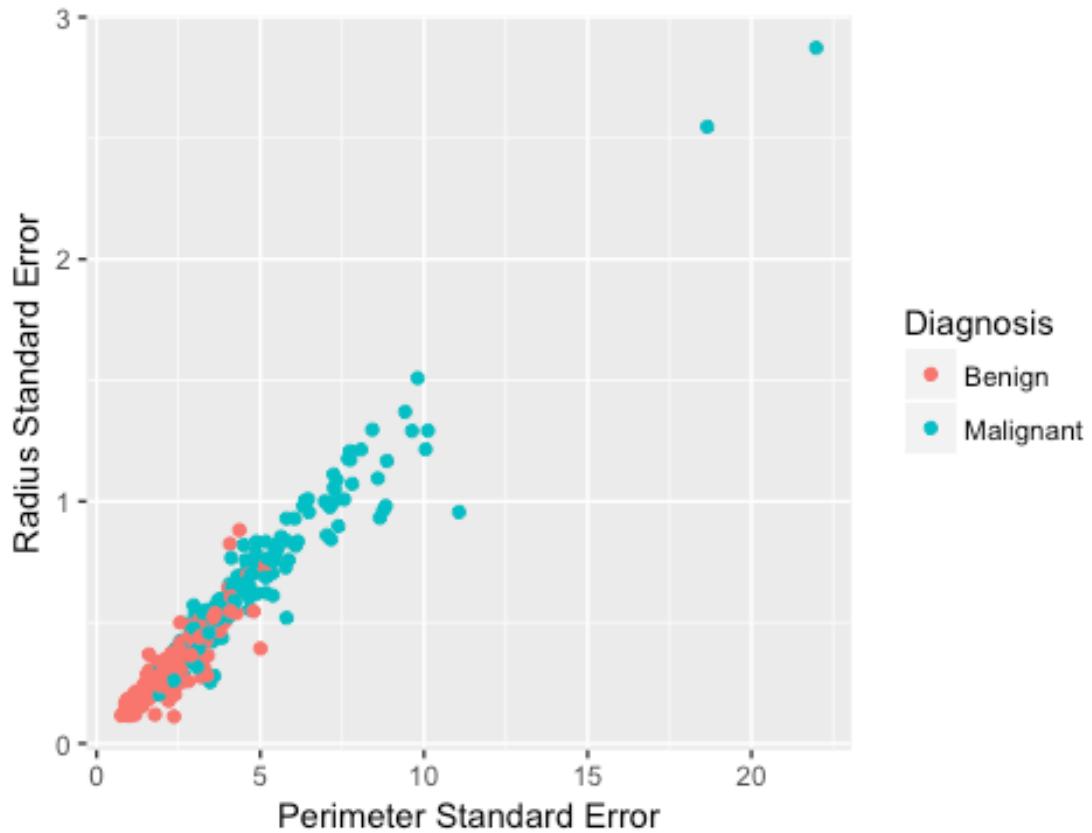


```
#fractal_dimension_se vs. texture_se
texture_fractal_dimension_se_scatter <- ggplot(data = cancer, aes(x = texture_se, y = fractal_dimension_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Texture Standard Error", y = "Fractal Dimension Standard Error", title = "Fractal Dimension Standard Error vs. Texture Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
texture_fractal_dimension_se_scatter
```



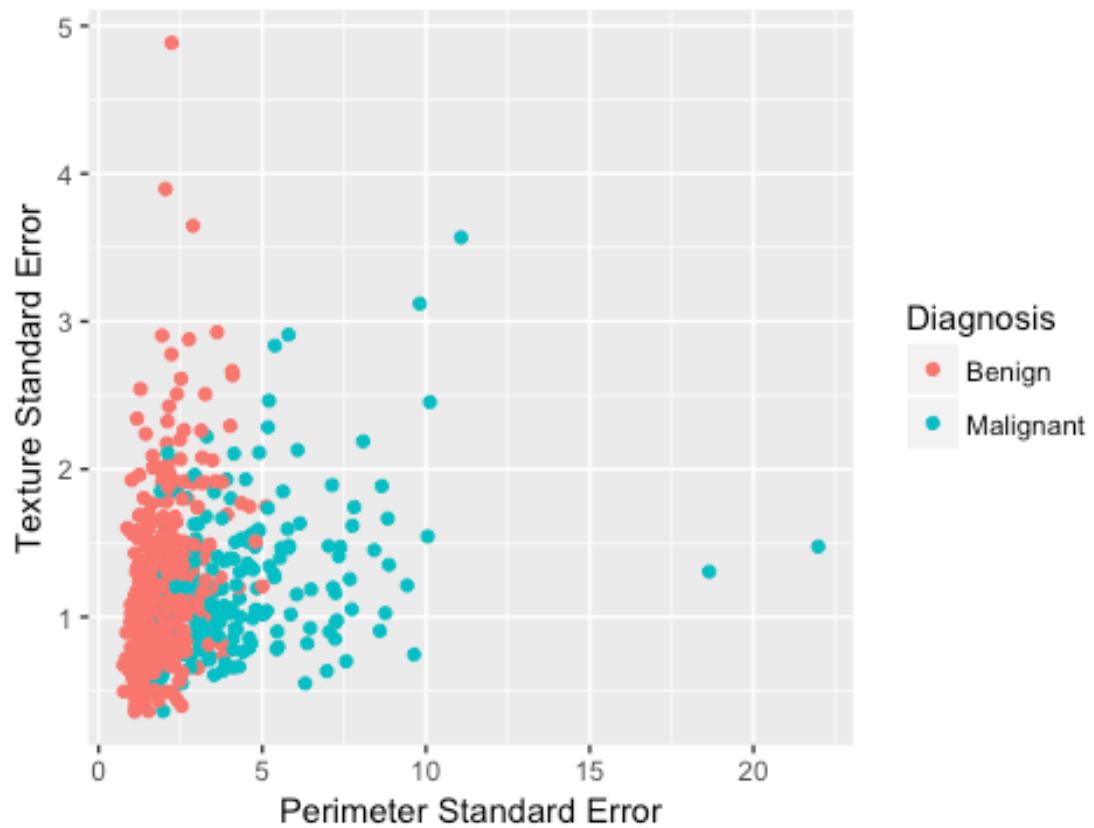
```
#radius_se vs. perimeter_se
perimeter_radius_se_scatter <- ggplot(data = cancer, aes(x = perimeter_se, y = radius_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Standard Error", y = "Radius Standard Error", title = "Radius Standard Error vs. Perimeter Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_radius_se_scatter
```

Radius Standard Error vs. Perimeter Standard Error



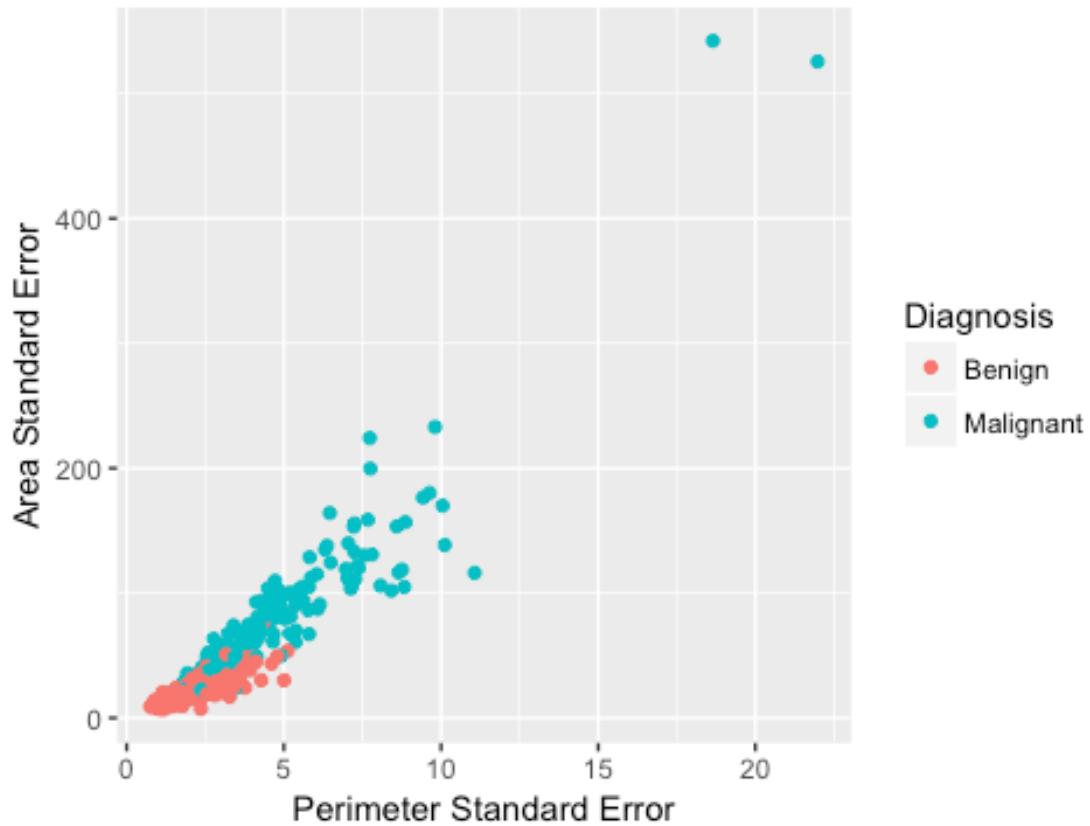
```
#texture_se vs. perimeter_se
perimeter_texture_se_scatter <- ggplot(data = cancer, aes(x = perimeter_se, y = texture_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Standard Error", y = "Texture Standard Error", title = "Texture Standard Error vs. Perimeter Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_texture_se_scatter
```

Texture Standard Error vs. Perimeter Standard Error



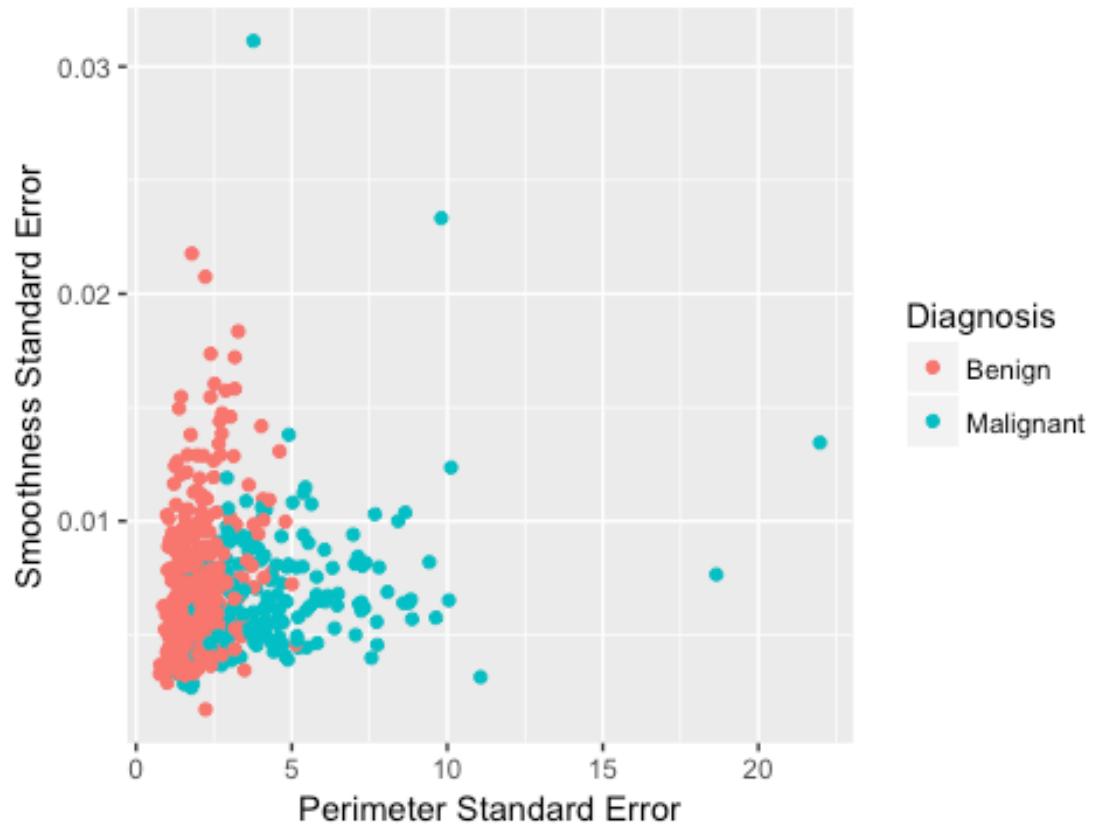
```
#area_se vs. perimeter_se
perimeter_area_se_scatter <- ggplot(data = cancer, aes(x = perimeter_se, y = area_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Standard Error", y = "Area Standard Error", title = "Area Standard Error vs. Perimeter Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_area_se_scatter
```

Area Standard Error vs. Perimeter Standard Error



```
#smoothness_se vs. perimeter_se
perimeter_smoothness_se_scatter <- ggplot(data = cancer, aes(x = perimeter_se,
, y = smoothness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Standard Error", y = "Smoothness Standard Error", title = "Smoothness Standard Error vs. Perimeter Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_smoothness_se_scatter
```

Smoothness Standard Error vs. Perimeter Standard Error



```
#compactness_se vs. perimeter_se
perimeter_compactness_se_scatter <- ggplot(data = cancer, aes(x = perimeter_se, y = compactness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Standard Error", y = "Compactness Standard Error", title = "Compactness Standard Error vs. Perimeter Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_compactness_se_scatter
```

Compactness Standard Error vs. Perimeter Standard Error



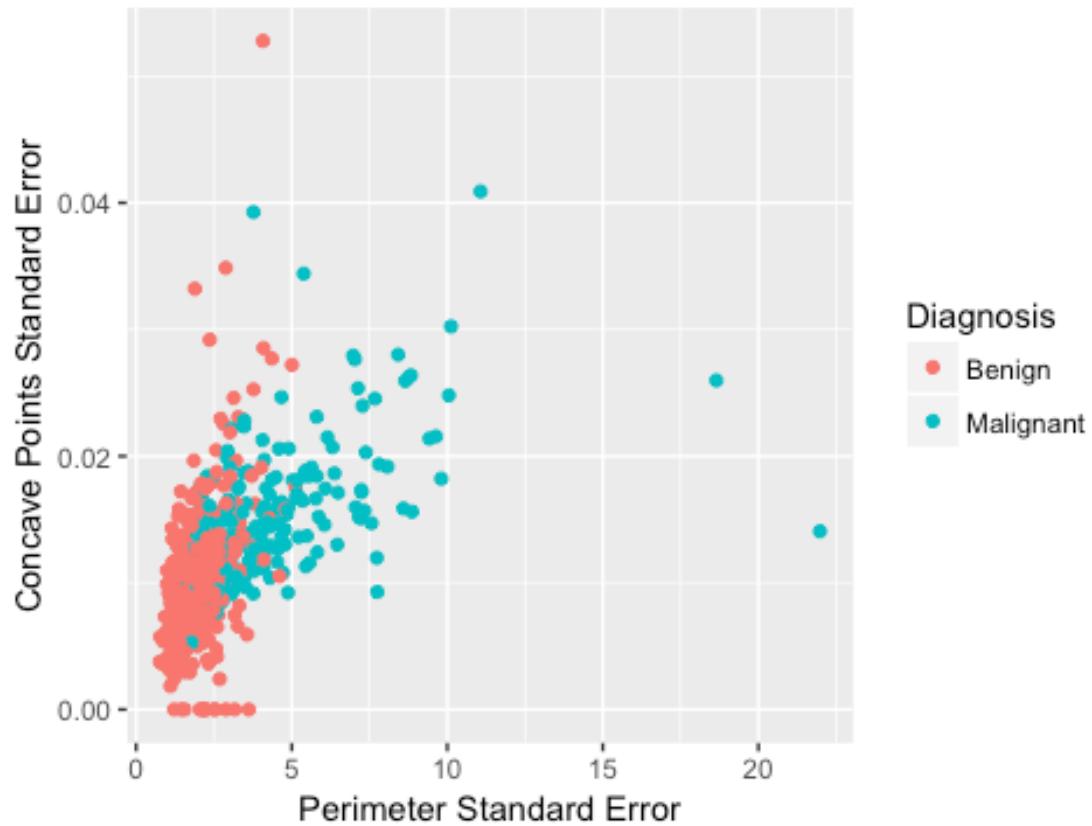
```
#concavity_se vs. perimeter_se
perimeter_concavity_se_scatter <- ggplot(data = cancer, aes(x = perimeter_se,
y = concavity_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Standard Error", y = "Concavity Standard
Error", title = "Concavity Standard Error vs. Perimeter Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_concavity_se_scatter
```

Concavity Standard Error vs. Perimeter Standard Error



```
#concave.points_se vs. perimeter_se
perimeter_concave.points_se_scatter <- ggplot(data = cancer, aes(x = perimeter_se, y = concave.points_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Standard Error", y = "Concave Points Standard Error", title = "Concave Points Standard Error vs. Perimeter Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_concave.points_se_scatter
```

Concave Points Standard Error vs. Perimeter Standard Error



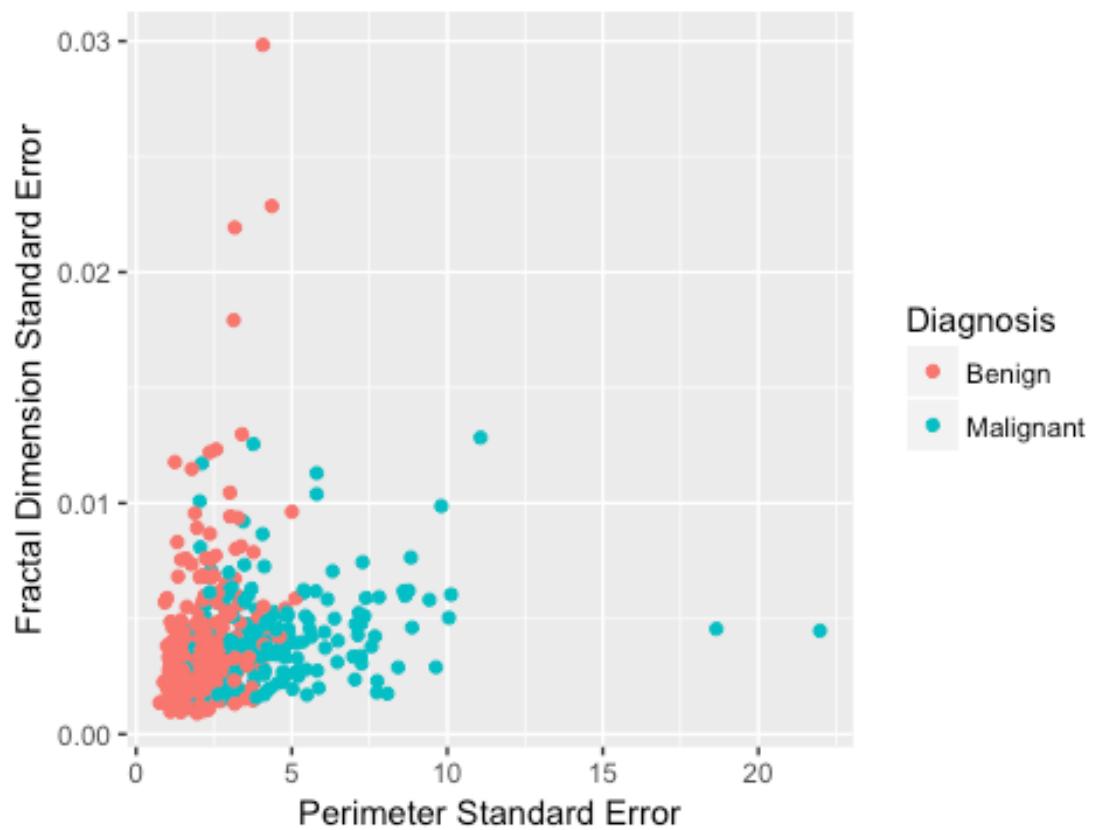
```
#symmetry_se vs. perimeter_se
perimeter_symmetry_se_scatter <- ggplot(data = cancer, aes(x = perimeter_se,
y = symmetry_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Standard Error", y = "Symmetry Standard
Error", title = "Symmetry Standard Error vs. Perimeter Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_symmetry_se_scatter
```

Symmetry Standard Error vs. Perimeter Standard Error



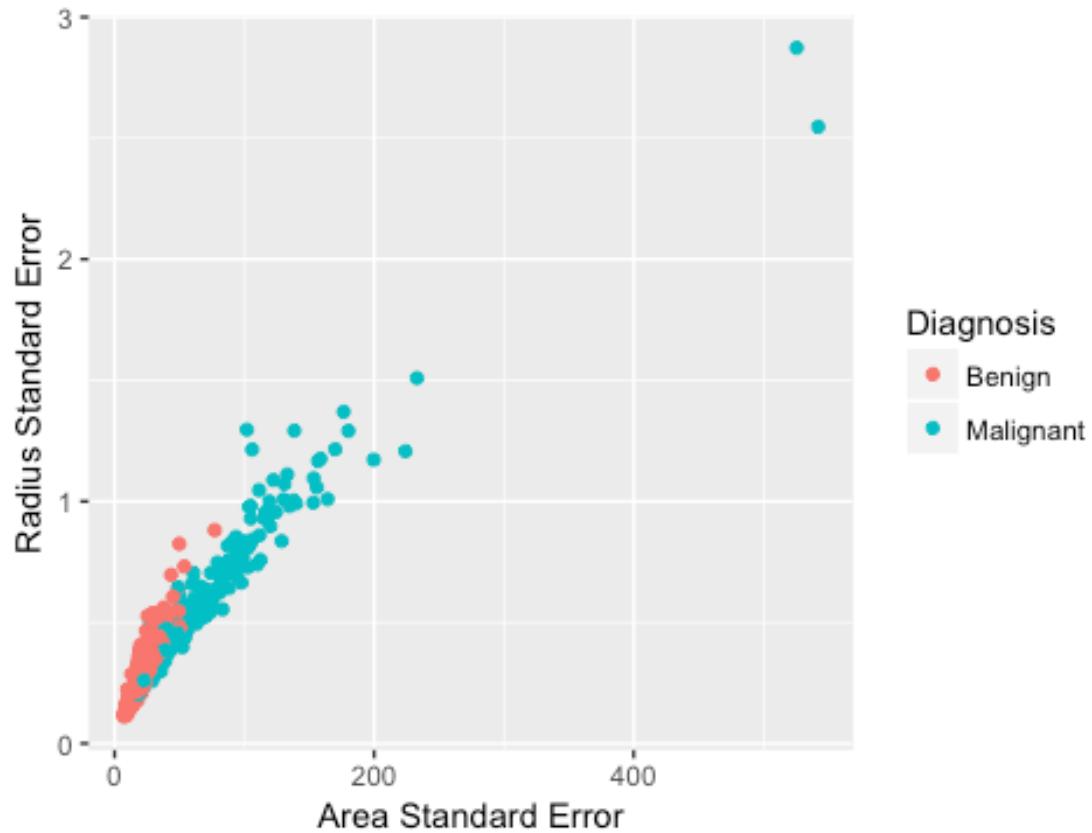
```
#fractal_dimension_se vs. perimeter_se
perimeter_fractal_dimension_se_scatter <- ggplot(data = cancer, aes(x = perimeter_se, y = fractal_dimension_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Perimeter Standard Error", y = "Fractal Dimension Standard Error", title = "Fractal Dimension Standard Error vs. Perimeter Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
perimeter_fractal_dimension_se_scatter
```

Fractal Dimension Standard Error vs. Perimeter Standard Error



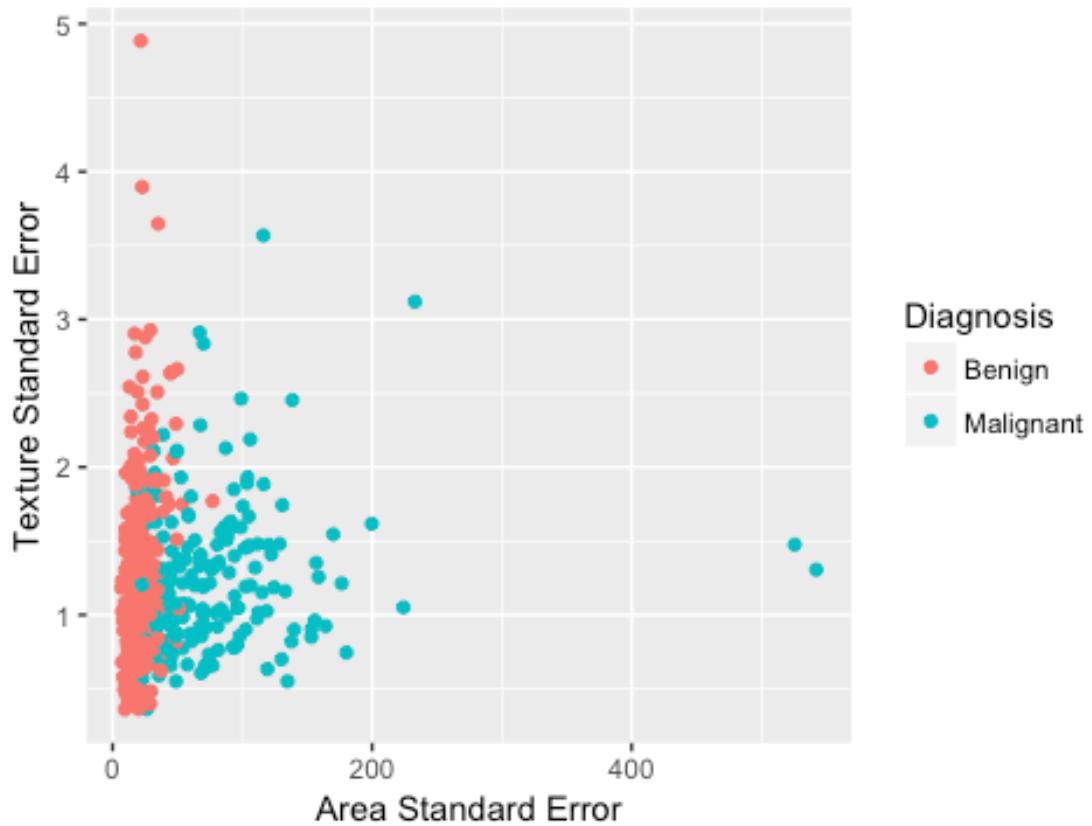
```
#radius_se vs. area_se
area_radius_se_scatter <- ggplot(data = cancer, aes(x = area_se, y = radius_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Standard Error", y = "Radius Standard Error",
  title = "Radius Standard Error vs. Area Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_radius_se_scatter
```

Radius Standard Error vs. Area Standard Error



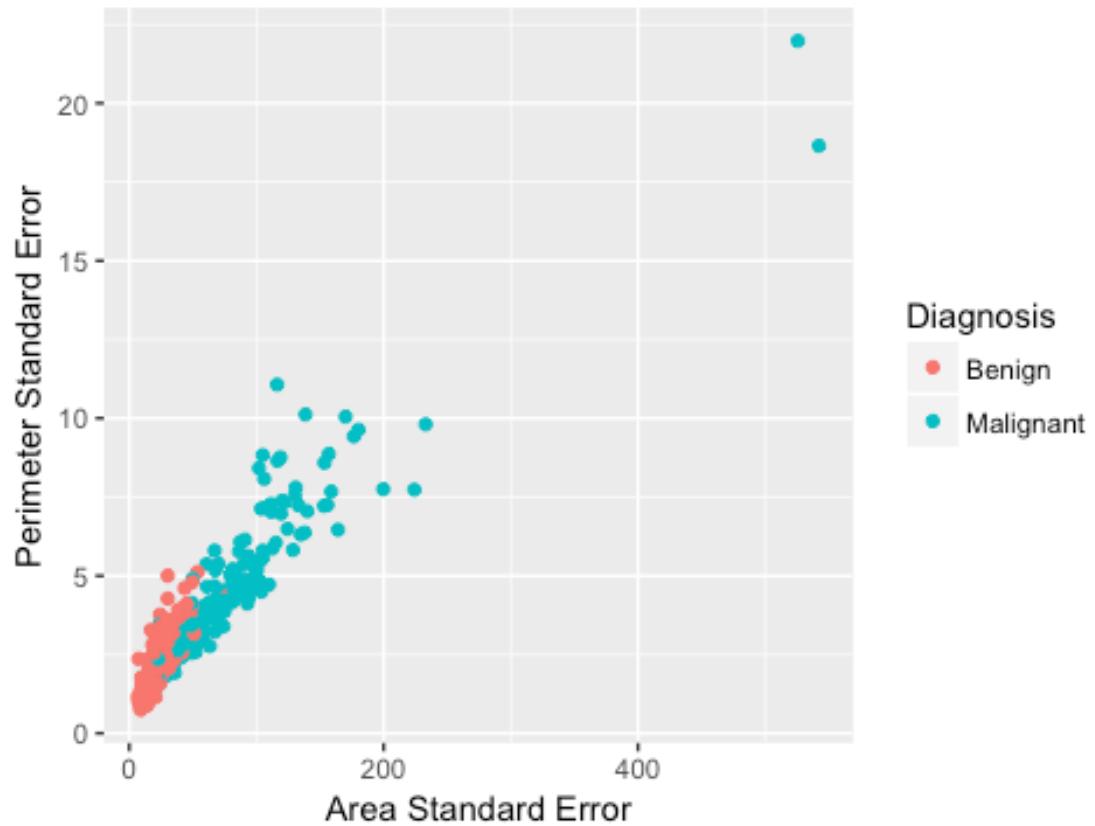
```
#texture_se vs. area_se
area_texture_se_scatter <- ggplot(data = cancer, aes(x = area_se, y = texture_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Standard Error", y = "Texture Standard Error",
, title = "Texture Standard Error vs. Area Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_texture_se_scatter
```

Texture Standard Error vs. Area Standard Error



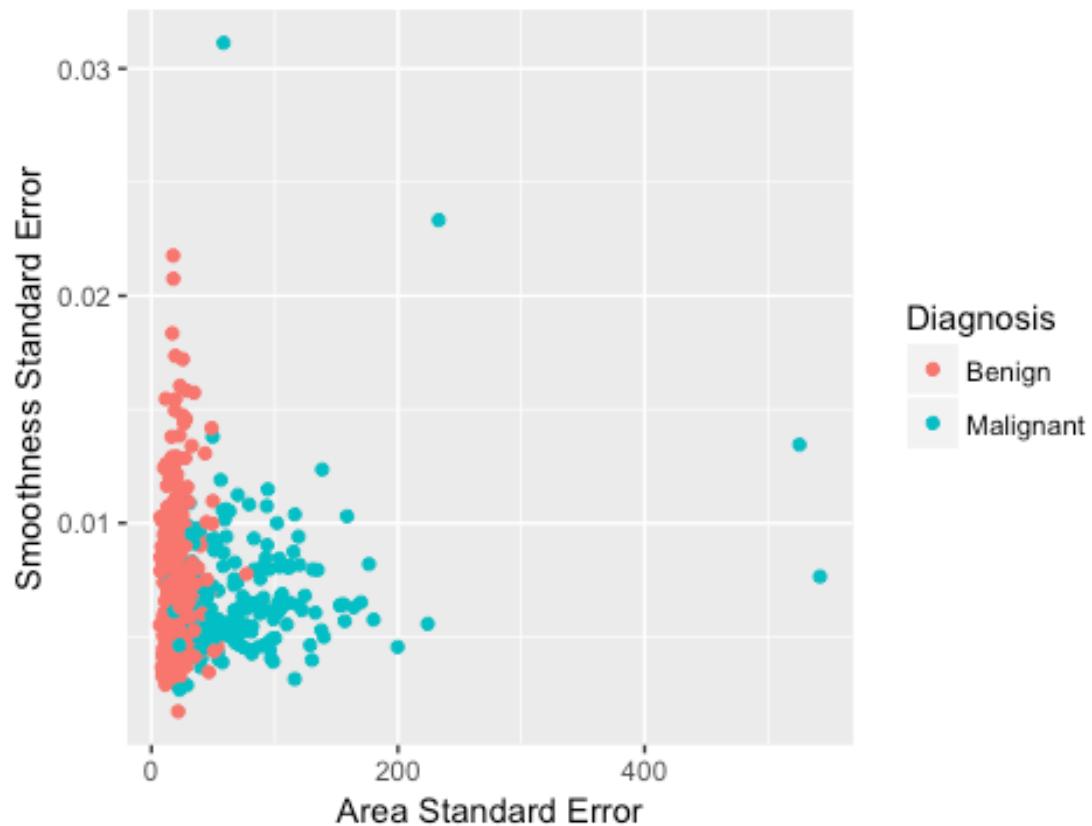
```
#perimeter_se vs. area_se
area_perimeter_se_scatter <- ggplot(data = cancer, aes(x = area_se, y = perimeter_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Standard Error", y = "Perimeter Standard Error", title = "Perimeter Standard Error vs. Area Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_perimeter_se_scatter
```

Perimeter Standard Error vs. Area Standard Error



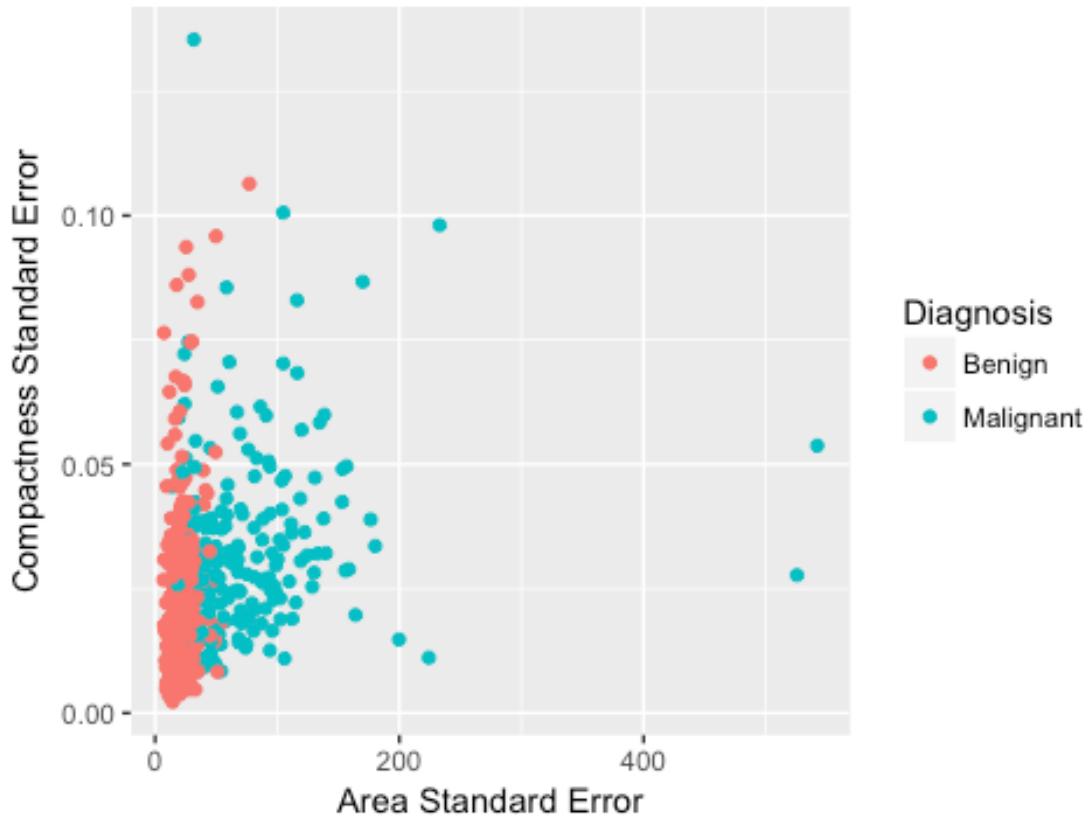
```
#smoothness_se vs. area_se
area_smoothness_se_scatter <- ggplot(data = cancer, aes(x = area_se, y = smoothness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Standard Error", y = "Smoothness Standard Error", title = "Smoothness Standard Error vs. Area Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_smoothness_se_scatter
```

Smoothness Standard Error vs. Area Standard Error



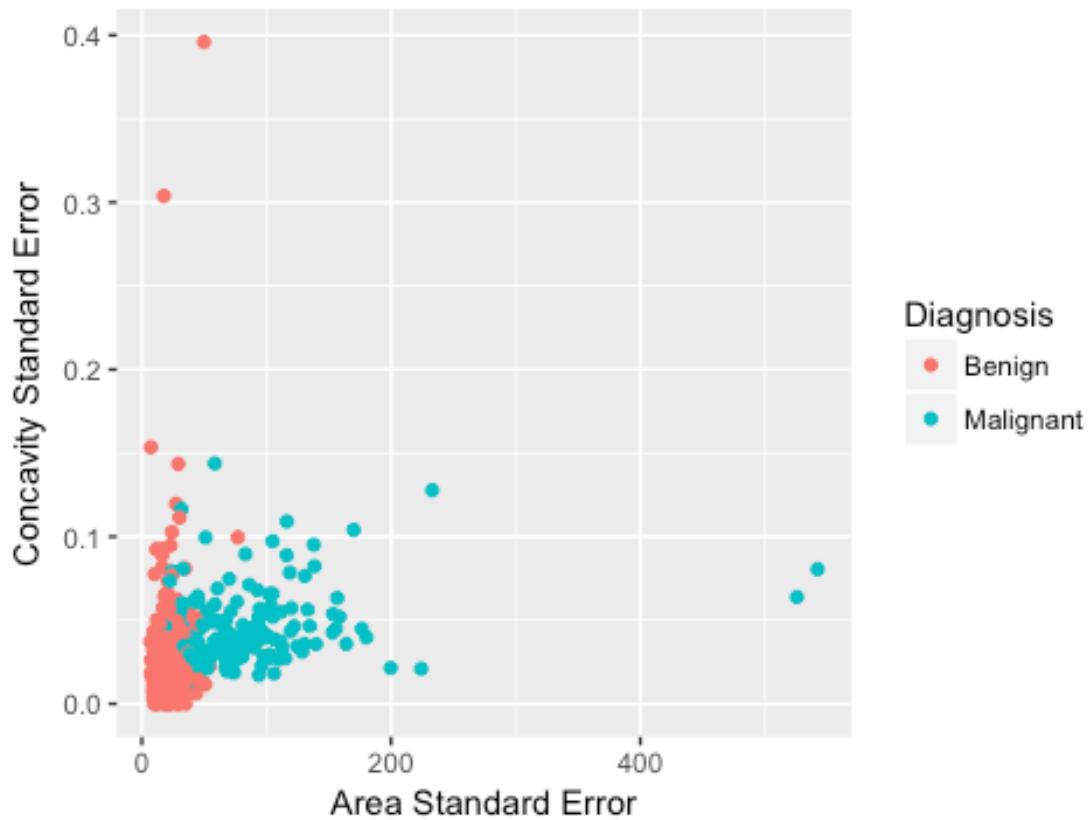
```
#compactness_se vs. area_se
area_compactness_se_scatter <- ggplot(data = cancer, aes(x = area_se, y = compactness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Standard Error", y = "Compactness Standard Error", title = "Compactness Standard Error vs. Area Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_compactness_se_scatter
```

Compactness Standard Error vs. Area Standard Error



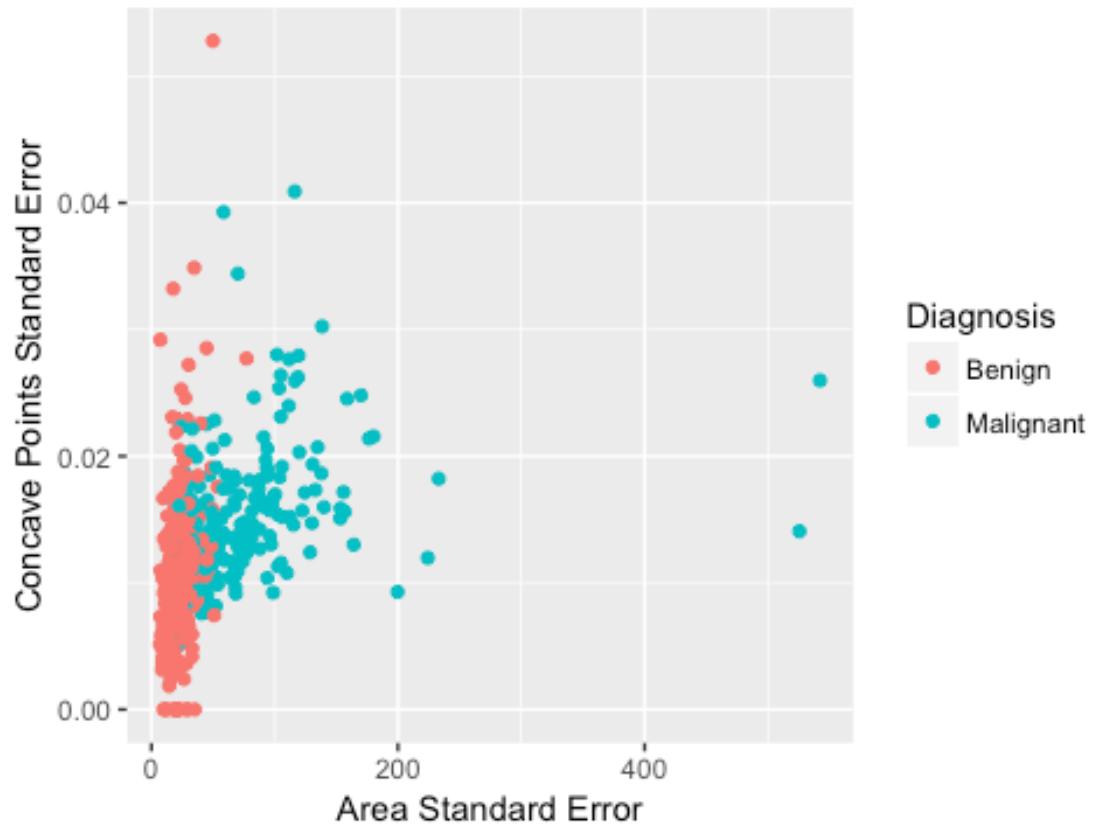
```
#concavity_se vs. area_se
area_concavity_se_scatter <- ggplot(data = cancer, aes(x = area_se, y = concavity_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Standard Error", y = "Concavity Standard Error", title = "Concavity Standard Error vs. Area Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_concavity_se_scatter
```

Concavity Standard Error vs. Area Standard Error



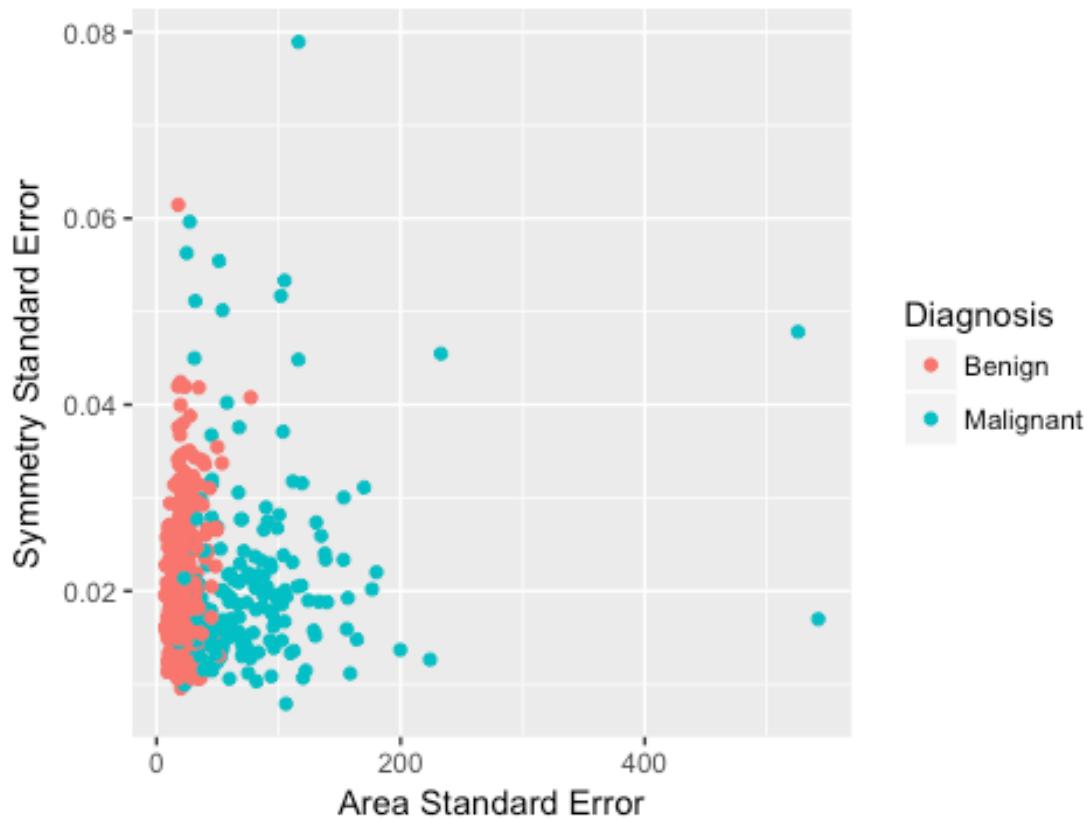
```
#concave.points_se vs. area_se
area_concave.points_se_scatter <- ggplot(data = cancer, aes(x = area_se, y = concave.points_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Standard Error", y = "Concave Points Standard Error", title = "Concave Points Standard Error vs. Area Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_concave.points_se_scatter
```

Concave Points Standard Error vs. Area Standard Error



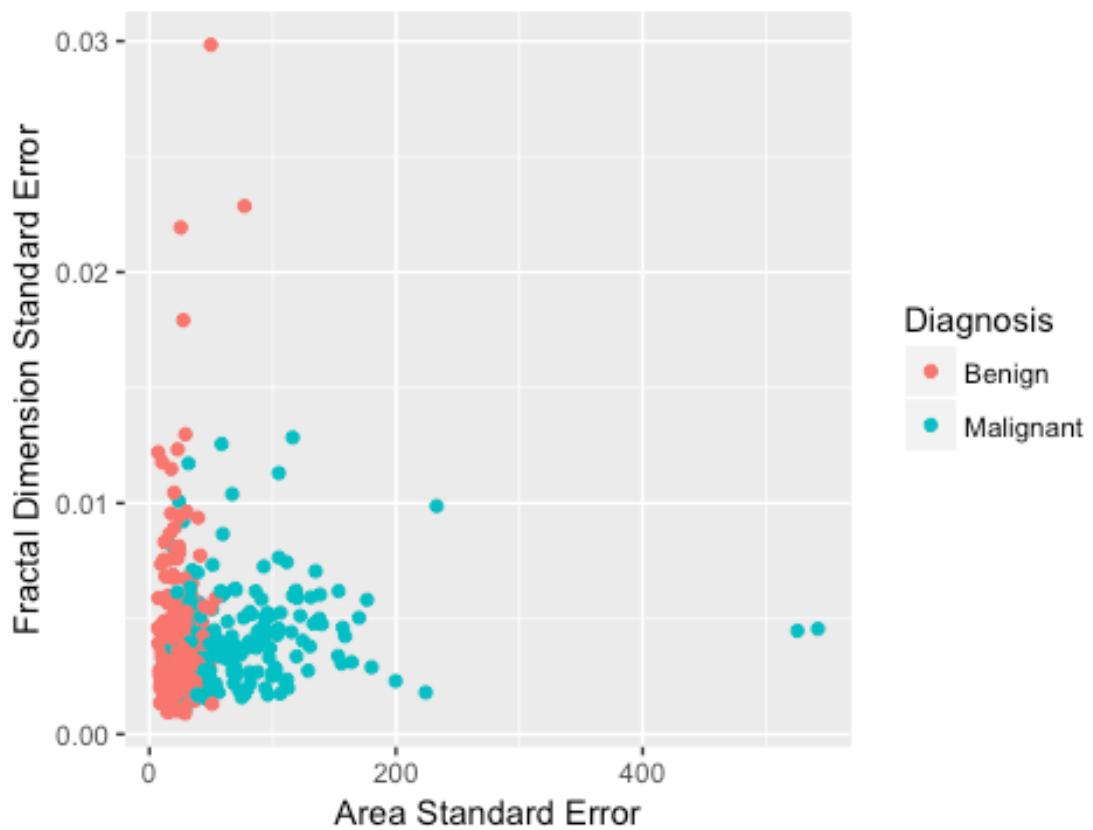
```
#symmetry_se vs. area_se
area_symmetry_se_scatter <- ggplot(data = cancer, aes(x = area_se, y = symmetry_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Standard Error", y = "Symmetry Standard Error",
  , title = "Symmetry Standard Error vs. Area Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_symmetry_se_scatter
```

Symmetry Standard Error vs. Area Standard Error



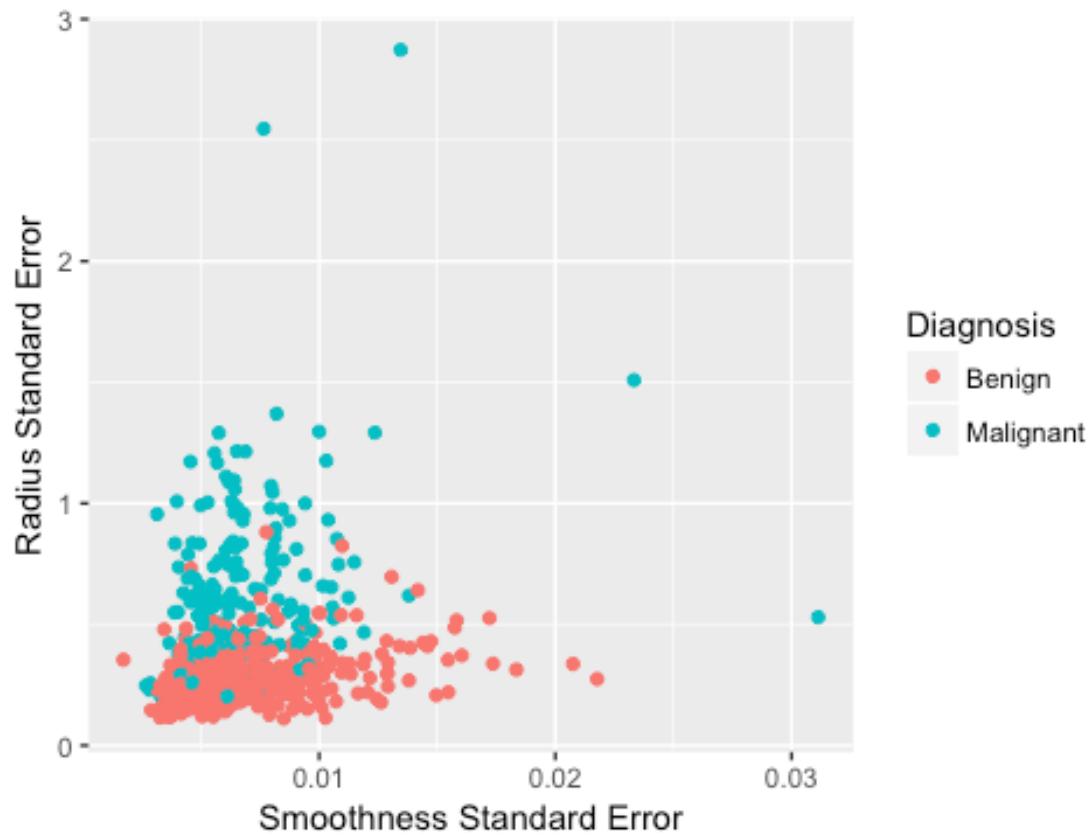
```
#fractal_dimension_se vs. area_se
area_fractal_dimension_se_scatter <- ggplot(data = cancer, aes(x = area_se, y = fractal_dimension_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Area Standard Error", y = "Fractal Dimension Standard Error", title = "Fractal Dimension Standard Error vs. Area Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
area_fractal_dimension_se_scatter
```

Fractal Dimension Standard Error vs. Area Standard E



```
#radius_se vs. smoothness_se
smoothness_radius_se_scatter <- ggplot(data = cancer, aes(x = smoothness_se,
y = radius_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Standard Error", y = "Radius Standard E
rror", title = "Radius Standard Error vs. Smoothness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_radius_se_scatter
```

Radius Standard Error vs. Smoothness Standard Error



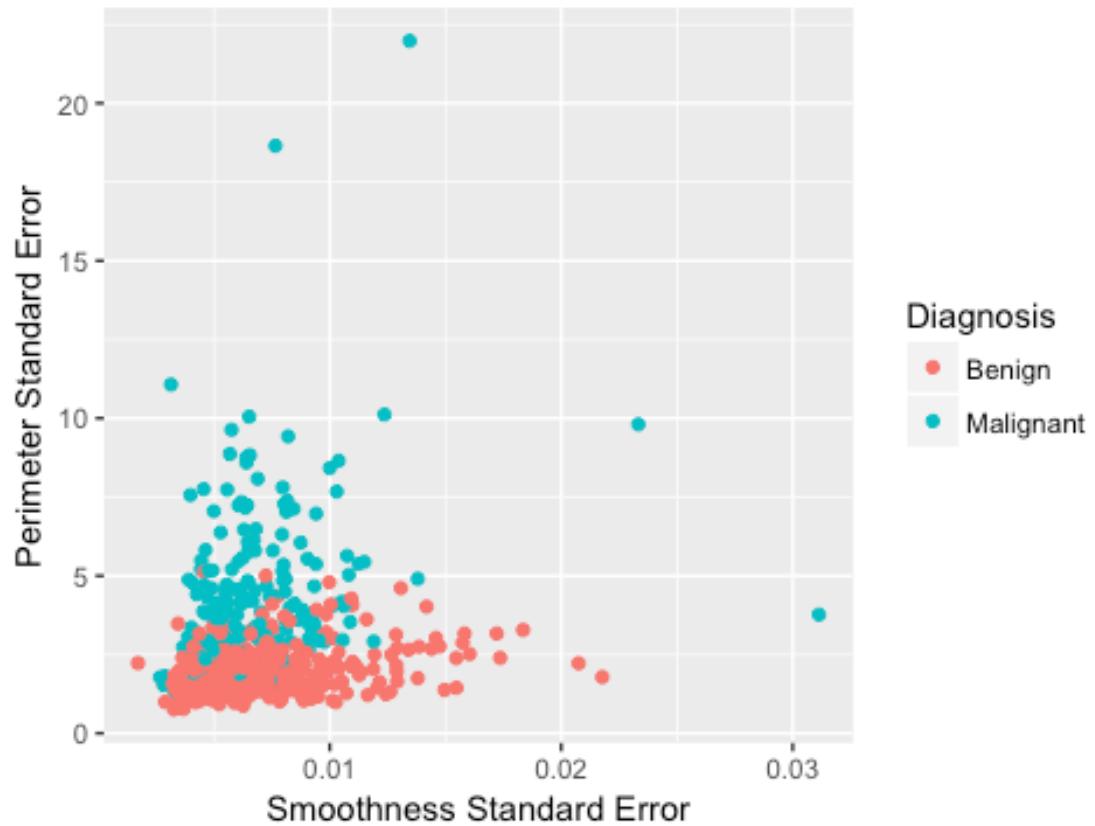
```
#texture_se vs. smoothness_se
smoothness_texture_se_scatter <- ggplot(data = cancer, aes(x = smoothness_se,
y = texture_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Standard Error", y = "Texture Standard
Error", title = "Texture Standard Error vs. Smoothness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_texture_se_scatter
```

Texture Standard Error vs. Smoothness Standard Error



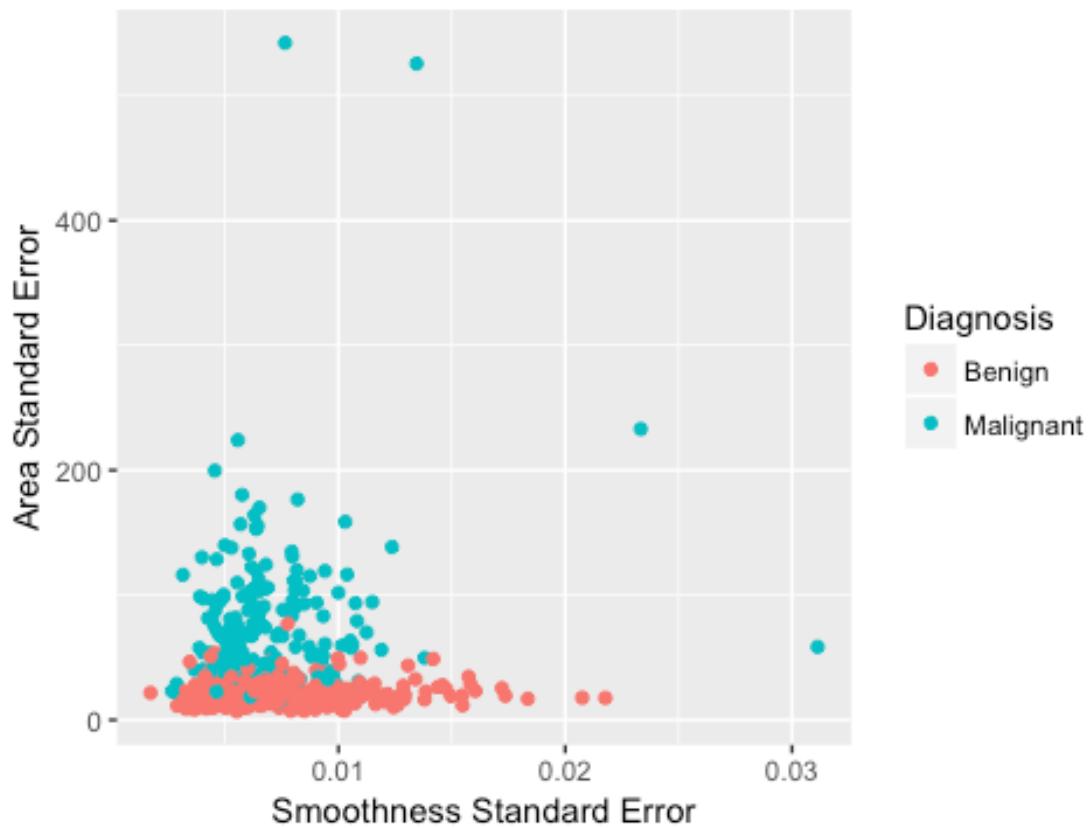
```
#perimeter_se vs. smoothness_se
smoothness_perimeter_se_scatter <- ggplot(data = cancer, aes(x = smoothness_se, y = perimeter_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Standard Error", y = "Perimeter Standard Error", title = "Perimeter Standard Error vs. Smoothness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_perimeter_se_scatter
```

Perimeter Standard Error vs. Smoothness Standard Error



```
#area_se vs. smoothness_se
smoothness_area_se_scatter <- ggplot(data = cancer, aes(x = smoothness_se, y = area_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Standard Error", y = "Area Standard Error", title = "Area Standard Error vs. Smoothness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_area_se_scatter
```

Area Standard Error vs. Smoothness Standard Error



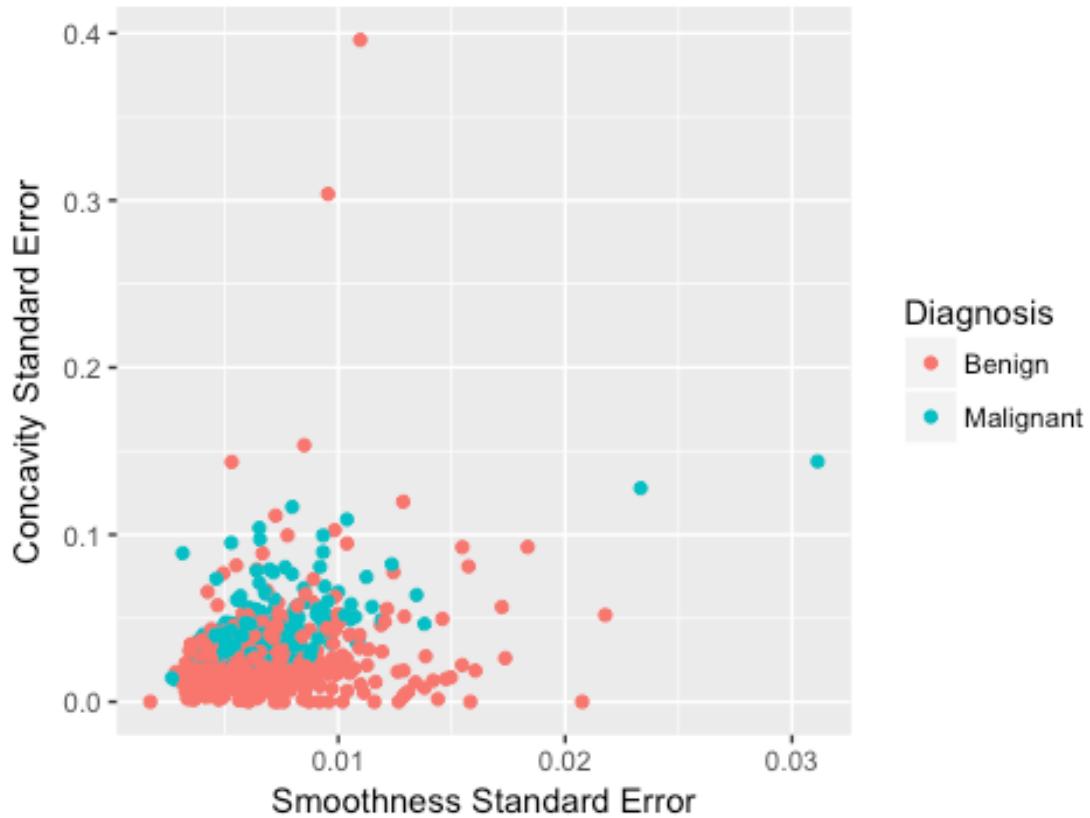
```
#compactness_se vs. smoothness_se
smoothness_compactness_se_scatter <- ggplot(data = cancer, aes(x = smoothness_se, y = compactness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Standard Error", y = "Compactness Standard Error", title = "Compactness Standard Error vs. Smoothness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_compactness_se_scatter
```

Compactness Standard Error vs. Smoothness Standard Error



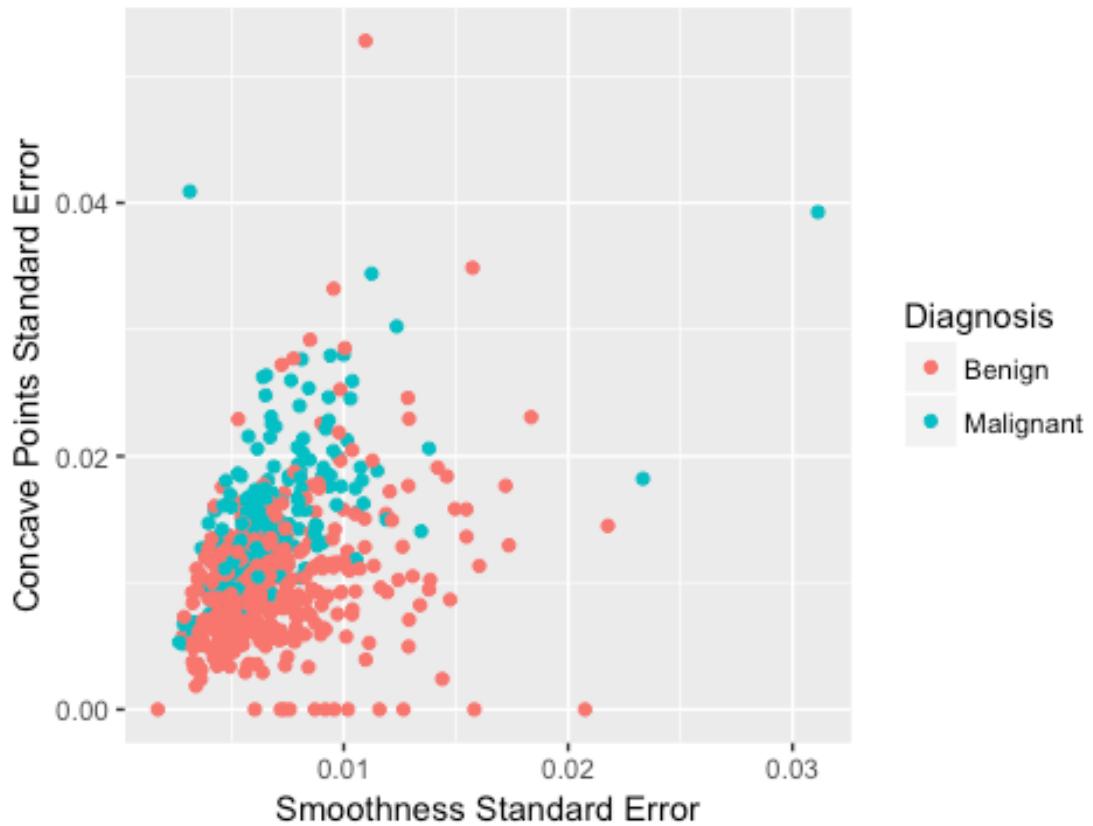
```
#concavity_se vs. smoothness_se
smoothness_concavity_se_scatter <- ggplot(data = cancer, aes(x = smoothness_se, y = concavity_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Standard Error", y = "Concavity Standard Error", title = "Concavity Standard Error vs. Smoothness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_concavity_se_scatter
```

Concavity Standard Error vs. Smoothness Standard Error



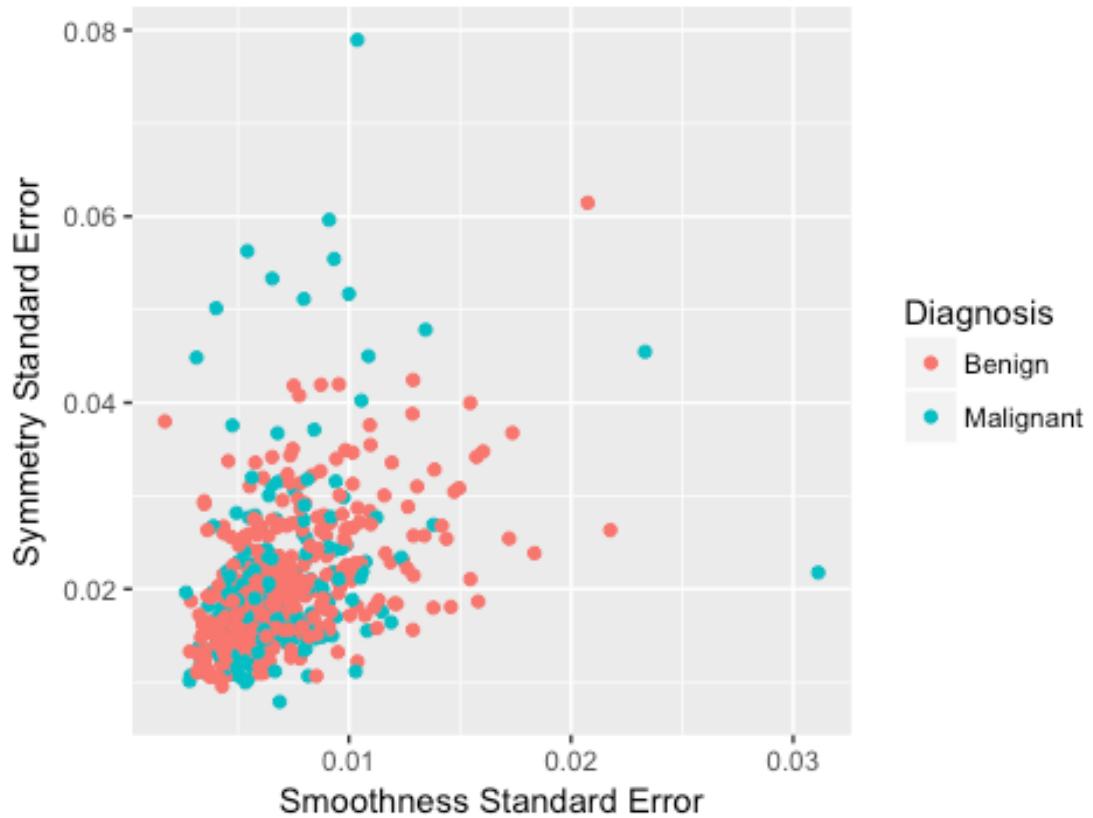
```
#concave.points_se vs. smoothness_se
smoothness_concave.points_se_scatter <- ggplot(data = cancer, aes(x = smoothness_se, y = concave.points_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Standard Error", y = "Concave Points Standard Error", title = "Concave Points Standard Error vs. Smoothness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_concave.points_se_scatter
```

Concave Points Standard Error vs. Smoothness Stand



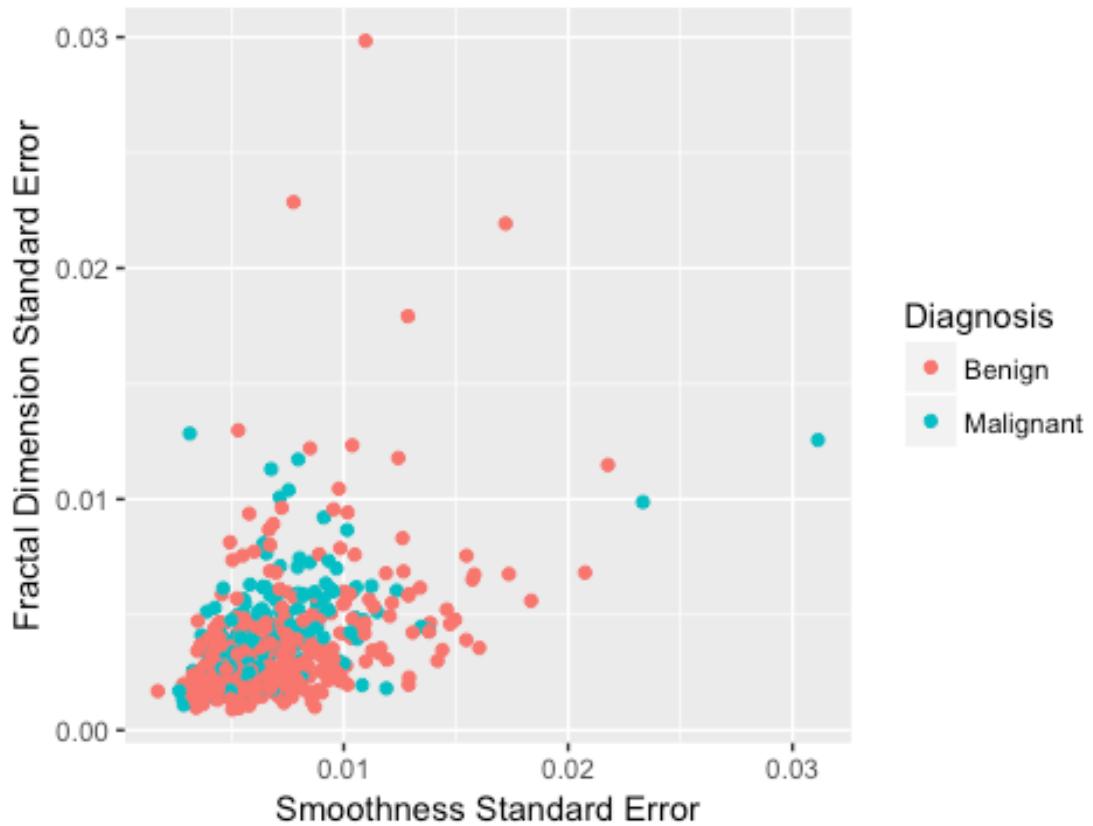
```
#symmetry_se vs. smoothness_se
smoothness_symmetry_se_scatter <- ggplot(data = cancer, aes(x = smoothness_se,
, y = symmetry_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Standard Error", y = "Symmetry Standard
Error", title = "Symmetry Standard Error vs. Smoothness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_symmetry_se_scatter
```

Symmetry Standard Error vs. Smoothness Standard Error



```
#fractal_dimension_se vs. smoothness_se
smoothness_fractal_dimension_se_scatter <- ggplot(data = cancer, aes(x = smoothness_se, y = fractal_dimension_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Smoothness Standard Error", y = "Fractal Dimension Standard Error", title = "Fractal Dimension Standard Error vs. Smoothness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
smoothness_fractal_dimension_se_scatter
```

Fractal Dimension Standard Error vs. Smoothness Standard Error



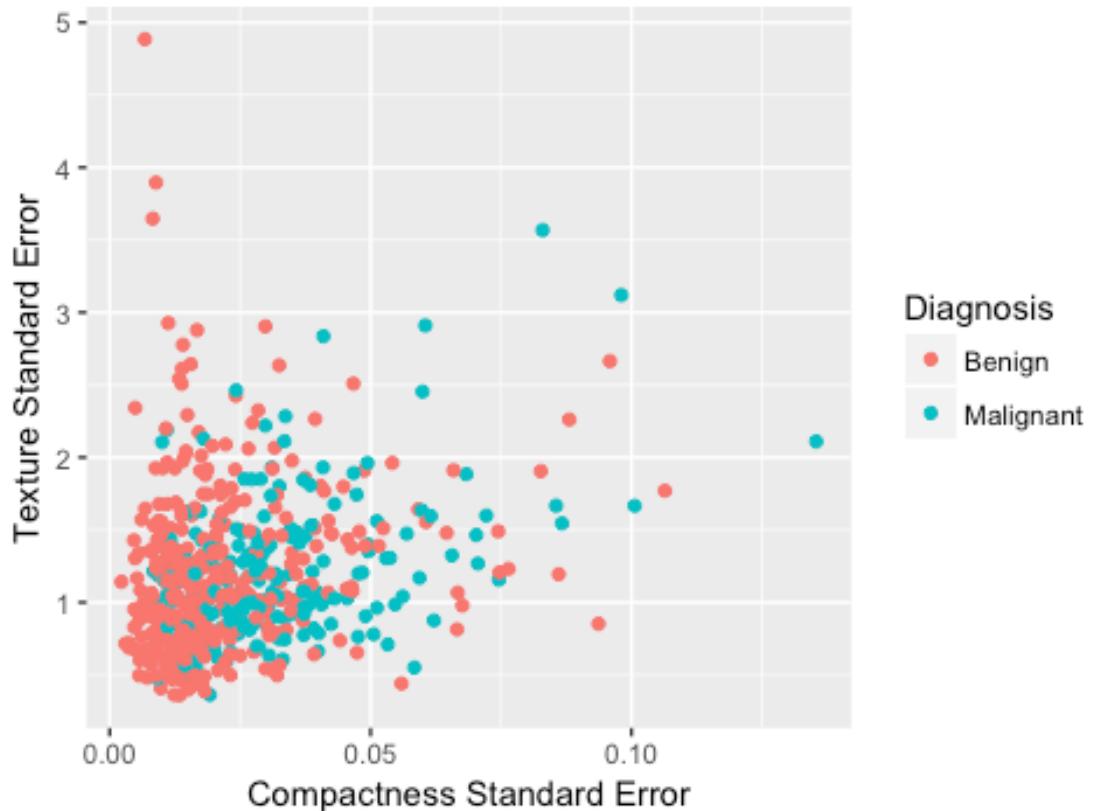
```
#radius_se vs. compactness_se
compactness_radius_se_scatter <- ggplot(data = cancer, aes(x = compactness_se,
, y = radius_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Standard Error", y = "Radius Standard
Error", title = "Radius Standard Error vs. Compactness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_radius_se_scatter
```

Radius Standard Error vs. Compactness Standard Error



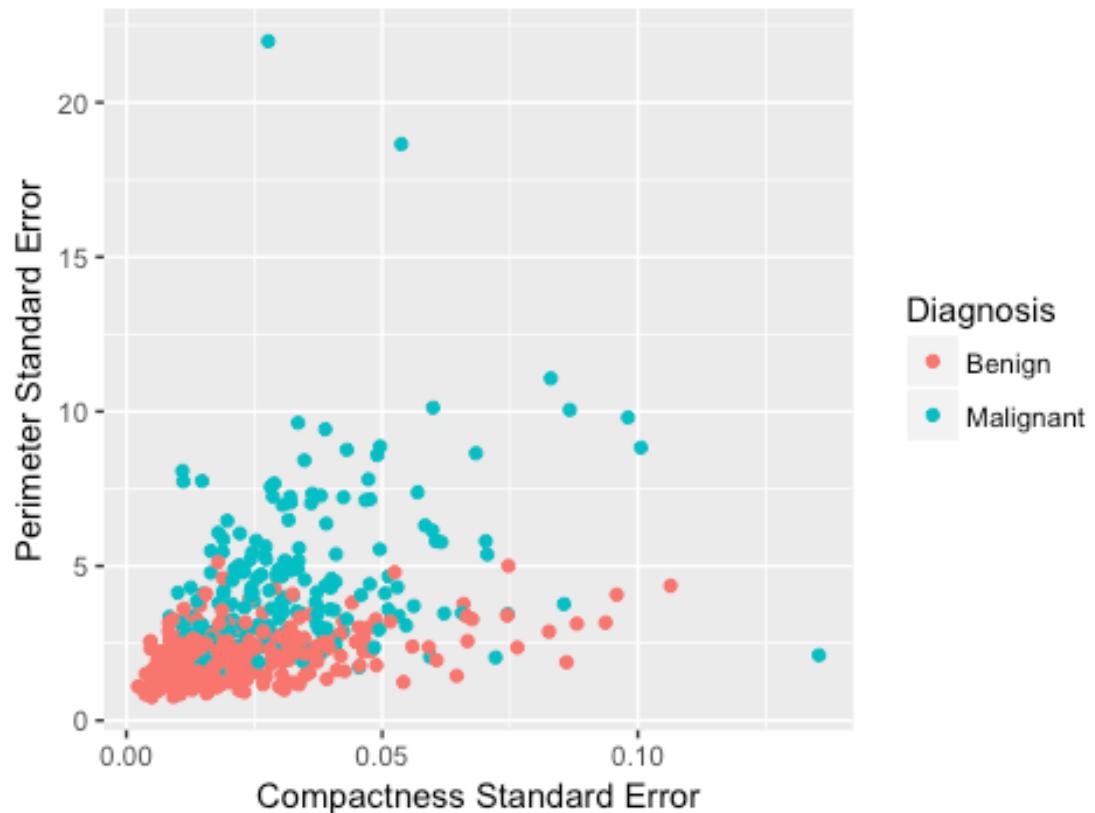
```
#texture_se vs. compactness_se
compactness_texture_se_scatter <- ggplot(data = cancer, aes(x = compactness_se, y = texture_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Standard Error", y = "Texture Standard Error", title = "Texture Standard Error vs. Compactness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_texture_se_scatter
```

Texture Standard Error vs. Compactness Standard Error



```
#perimeter_se vs. compactness_se
compactness_perimeter_se_scatter <- ggplot(data = cancer, aes(x = compactness_se, y = perimeter_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Standard Error", y = "Perimeter Standard Error", title = "Perimeter Standard Error vs. Compactness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_perimeter_se_scatter
```

Perimeter Standard Error vs. Compactness Standard Error



```
#area_se vs. compactness_se
compactness_area_se_scatter <- ggplot(data = cancer, aes(x = compactness_se,
y = area_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Standard Error", y = "Area Standard Error",
title = "Area Standard Error vs. Compactness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_area_se_scatter
```

Area Standard Error vs. Compactness Standard Error



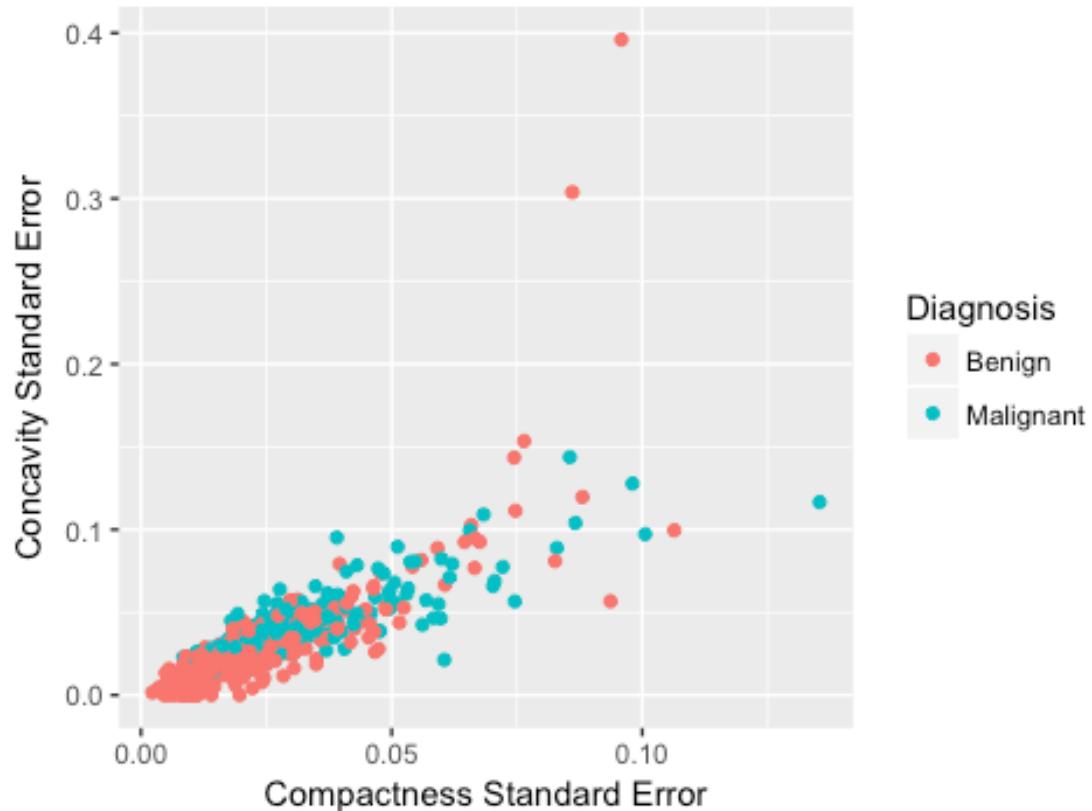
```
#smoothness_se vs. compactness_se
compactness_smoothness_se_scatter <- ggplot(data = cancer, aes(x = compactness_se, y = smoothness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Standard Error", y = "Smoothness Standard Error", title = "Smoothness Standard Error vs. Compactness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_smoothness_se_scatter
```

Smoothness Standard Error vs. Compactness Standard Error



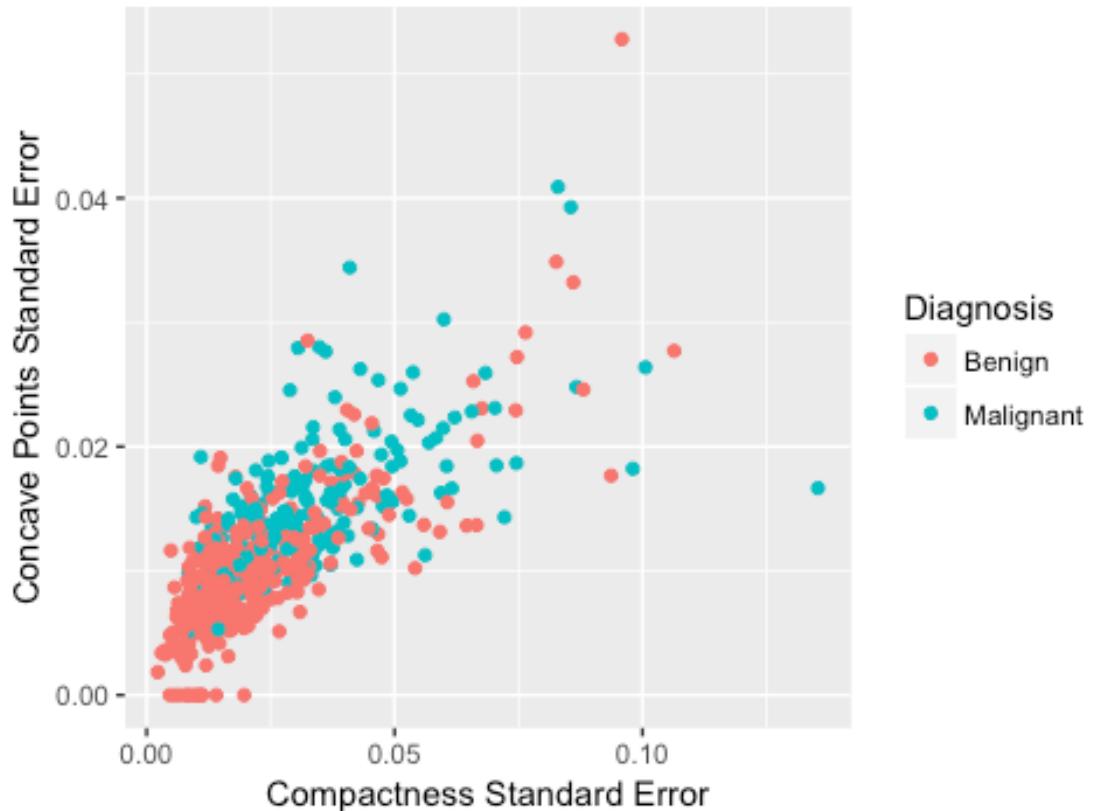
```
#concavity_se vs. compactness_se
compactness_concavity_se_scatter <- ggplot(data = cancer, aes(x = compactness_se, y = concavity_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Standard Error", y = "Concavity Standard Error", title = "Concavity Standard Error vs. Compactness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_concavity_se_scatter
```

Concavity Standard Error vs. Compactness Standard Error



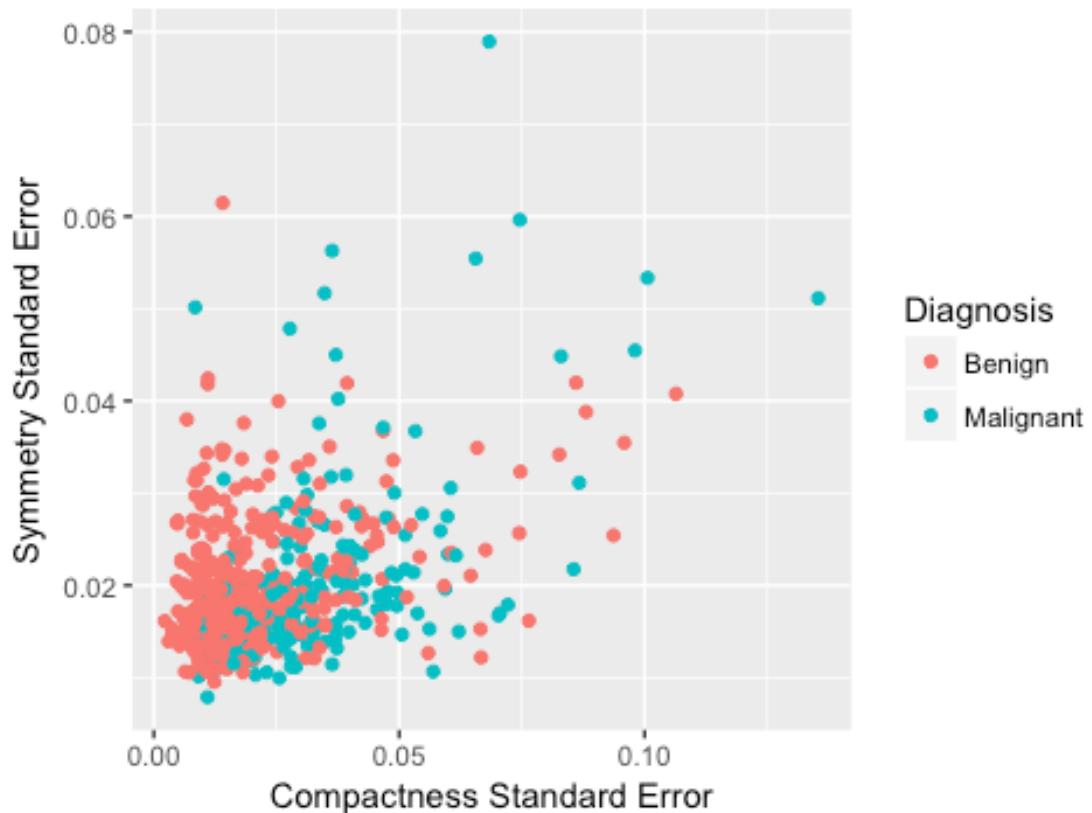
```
#concave.points_se vs. compactness_se
compactness_concave.points_se_scatter <- ggplot(data = cancer, aes(x = compactness_se, y = concave.points_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Standard Error", y = "Concave Points Standard Error", title = "Concave Points Standard Error vs. Compactness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_concave.points_se_scatter
```

Concave Points Standard Error vs. Compactness Standard Error

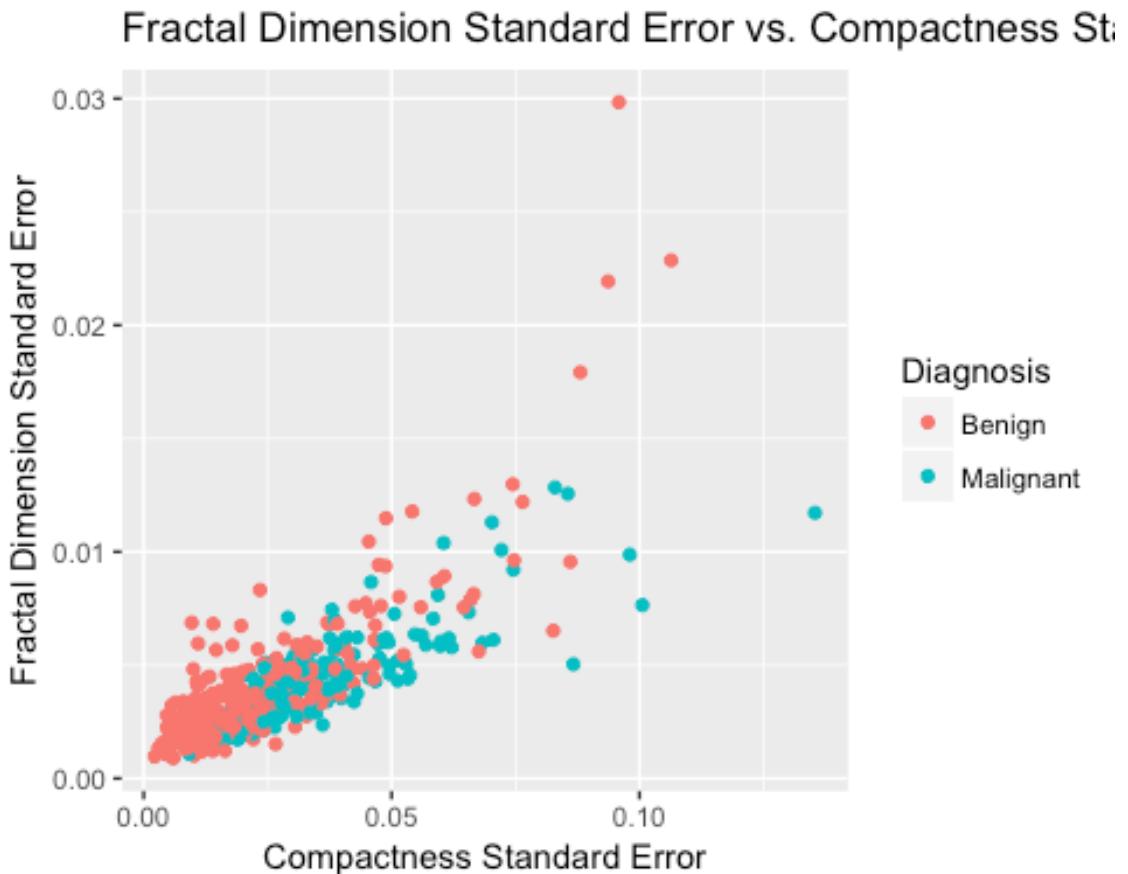


```
#symmetry_se vs. compactness_se
compactness_symmetry_se_scatter <- ggplot(data = cancer, aes(x = compactness_se, y = symmetry_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Standard Error", y = "Symmetry Standard Error", title = "Symmetry Standard Error vs. Compactness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_symmetry_se_scatter
```

Symmetry Standard Error vs. Compactness Standard Error

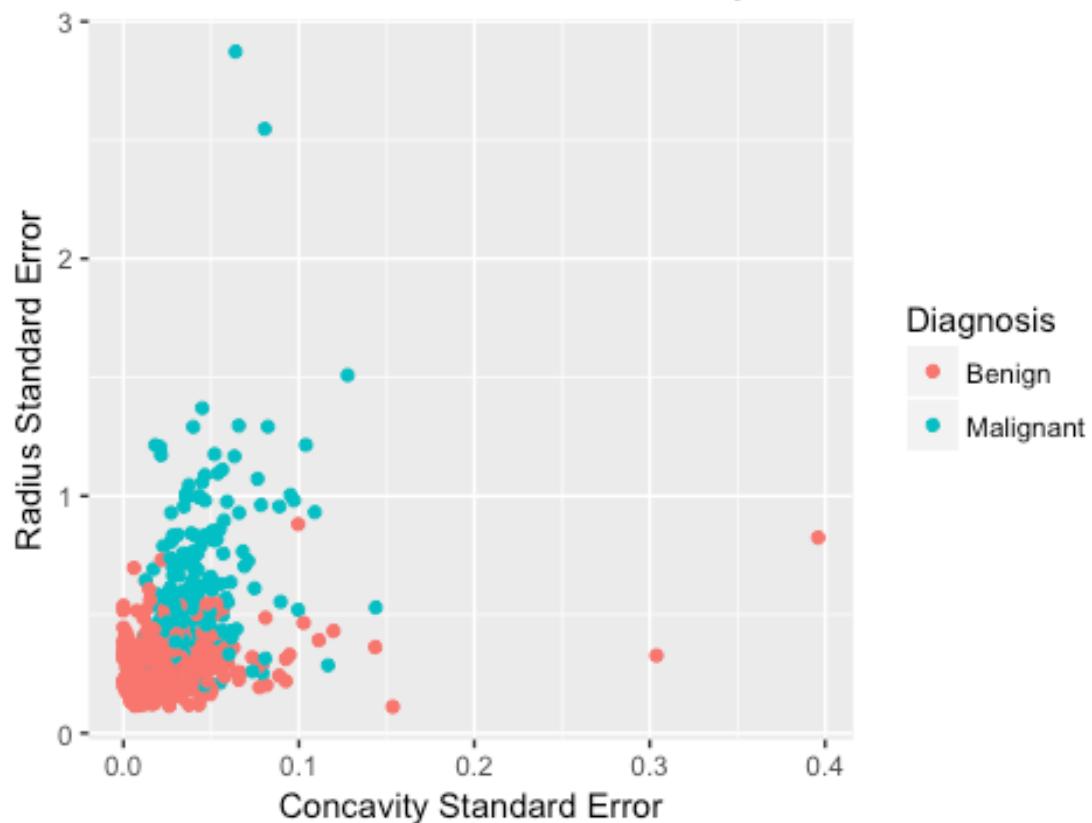


```
#fractal_dimension_se vs. compactness_se
compactness_fractal_dimension_se_scatter <- ggplot(data = cancer, aes(x = compactness_se, y = fractal_dimension_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Compactness Standard Error", y = "Fractal Dimension Standard Error", title = "Fractal Dimension Standard Error vs. Compactness Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
compactness_fractal_dimension_se_scatter
```



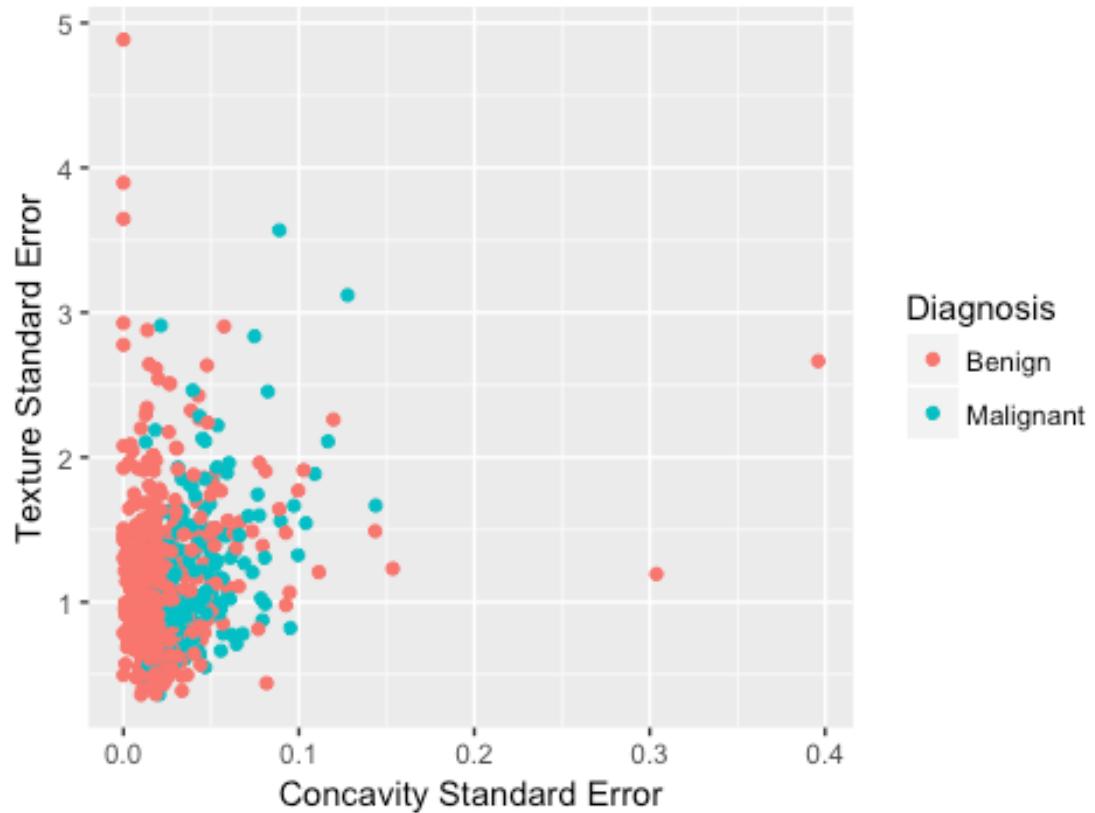
```
#radius_se vs. concavity_se
concavity_radius_se_scatter <- ggplot(data = cancer, aes(x = concavity_se, y = radius_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Standard Error", y = "Radius Standard Error", title = "Radius Standard Error vs. Concavity Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_radius_se_scatter
```

Radius Standard Error vs. Concavity Standard Error



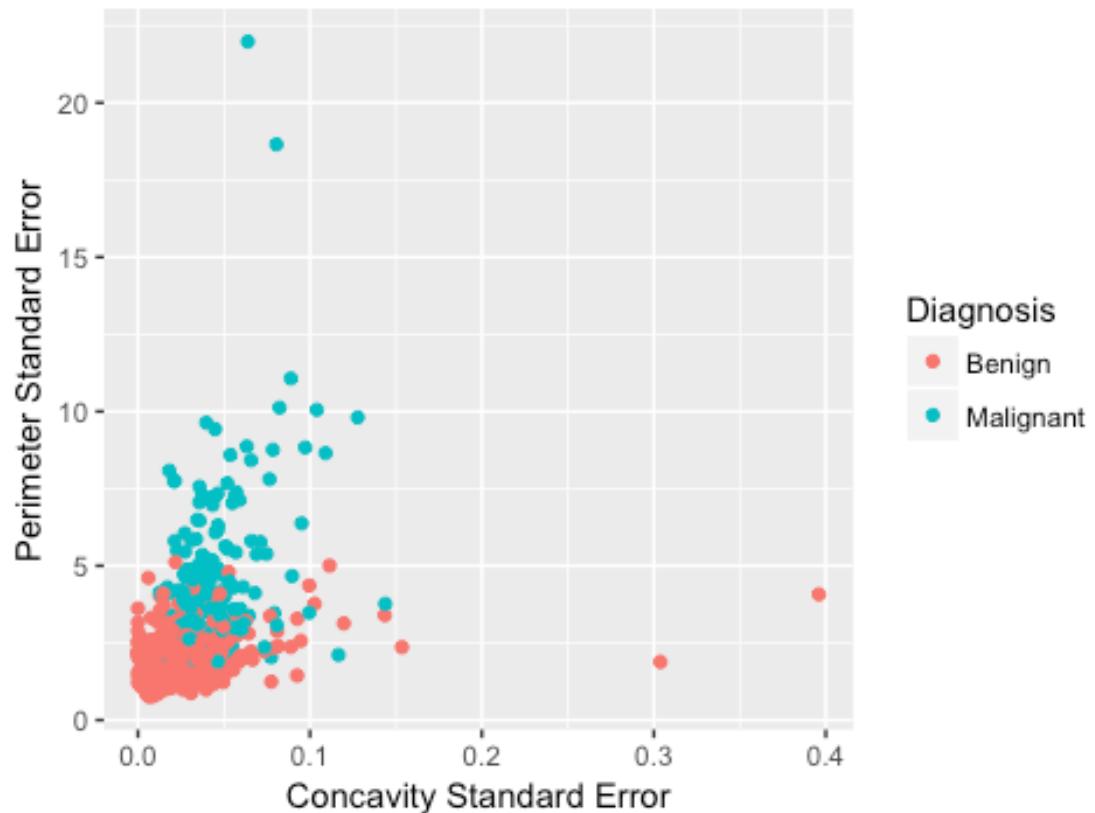
```
#texture_se vs. concavity_se
concavity_texture_se_scatter <- ggplot(data = cancer, aes(x = concavity_se, y = texture_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Standard Error", y = "Texture Standard Error", title = "Texture Standard Error vs. Concavity Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_texture_se_scatter
```

Texture Standard Error vs. Concavity Standard Error



```
#perimeter_se vs. concavity_se
concavity_perimeter_se_scatter <- ggplot(data = cancer, aes(x = concavity_se,
y = perimeter_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Standard Error", y = "Perimeter Standard
Error", title = "Perimeter Standard Error vs. Concavity Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_perimeter_se_scatter
```

Perimeter Standard Error vs. Concavity Standard Error



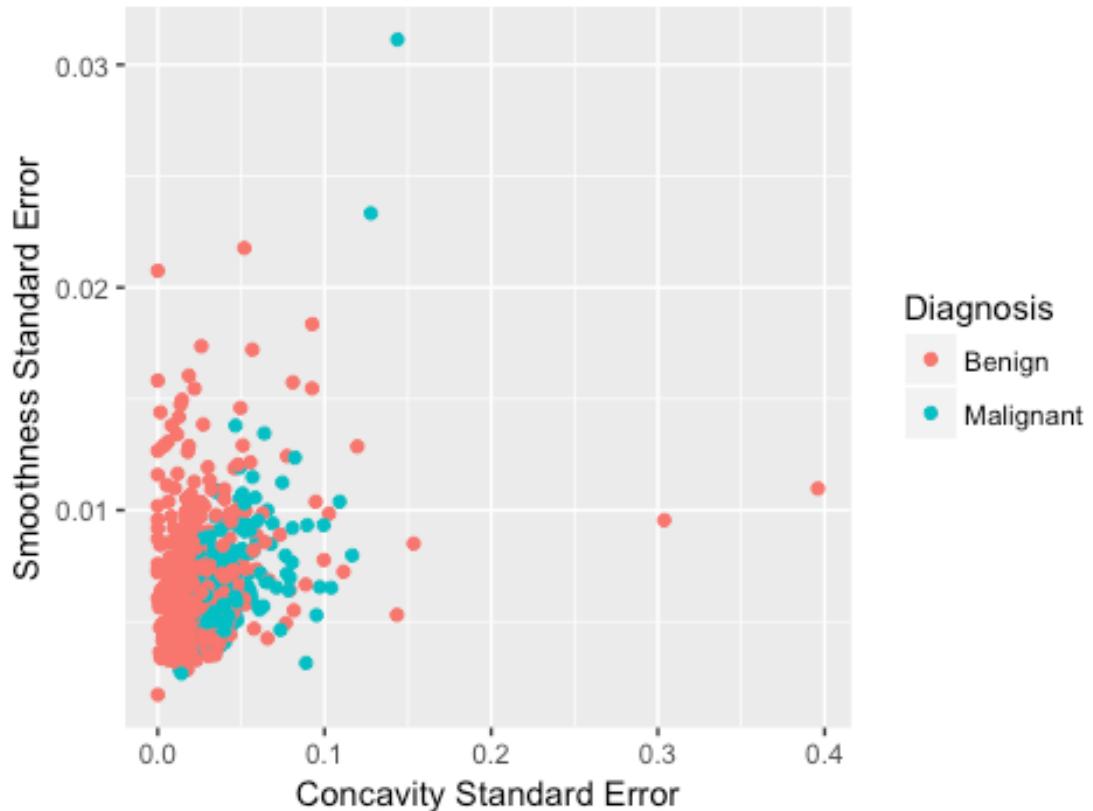
```
#area_se vs. concavity_se
concavity_area_se_scatter <- ggplot(data = cancer, aes(x = concavity_se, y = area_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Standard Error", y = "Area Standard Error", title = "Area Standard Error vs. Concavity Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_area_se_scatter
```

Area Standard Error vs. Concavity Standard Error



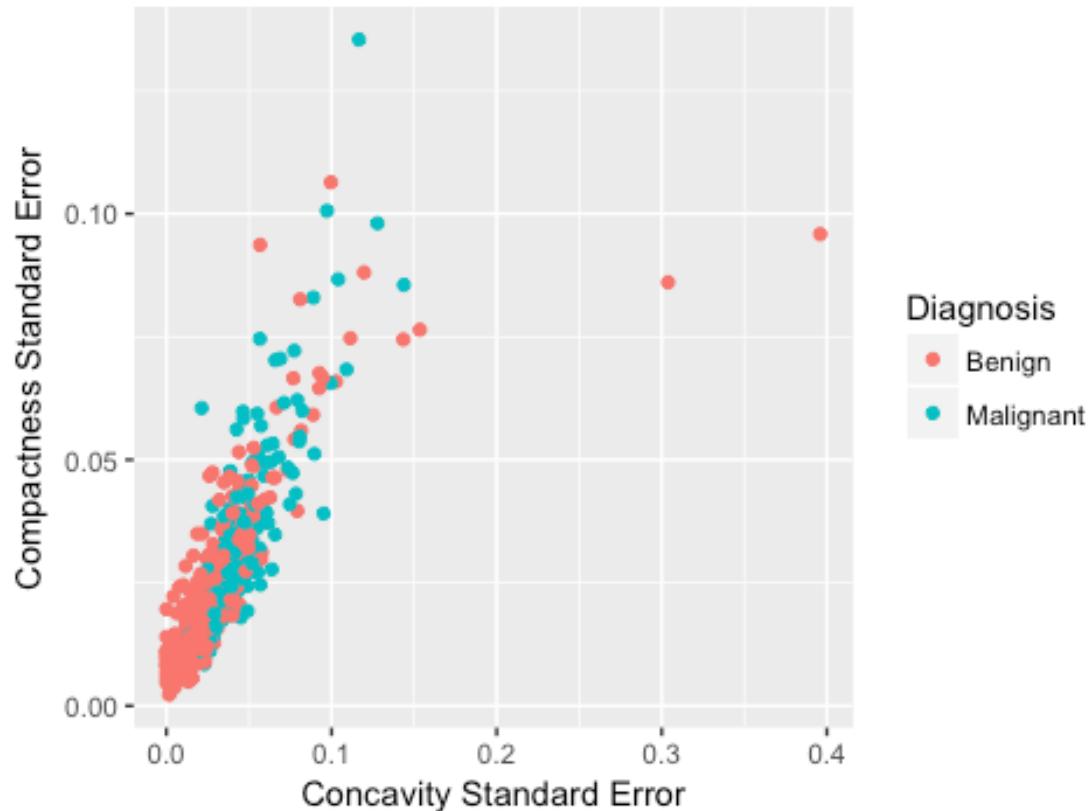
```
#smoothness_se vs. concavity_se
concavity_smoothness_se_scatter <- ggplot(data = cancer, aes(x = concavity_se,
, y = smoothness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Standard Error", y = "Smoothness Standard Error", title = "Smoothness Standard Error vs. Concavity Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_smoothness_se_scatter
```

Smoothness Standard Error vs. Concavity Standard Er



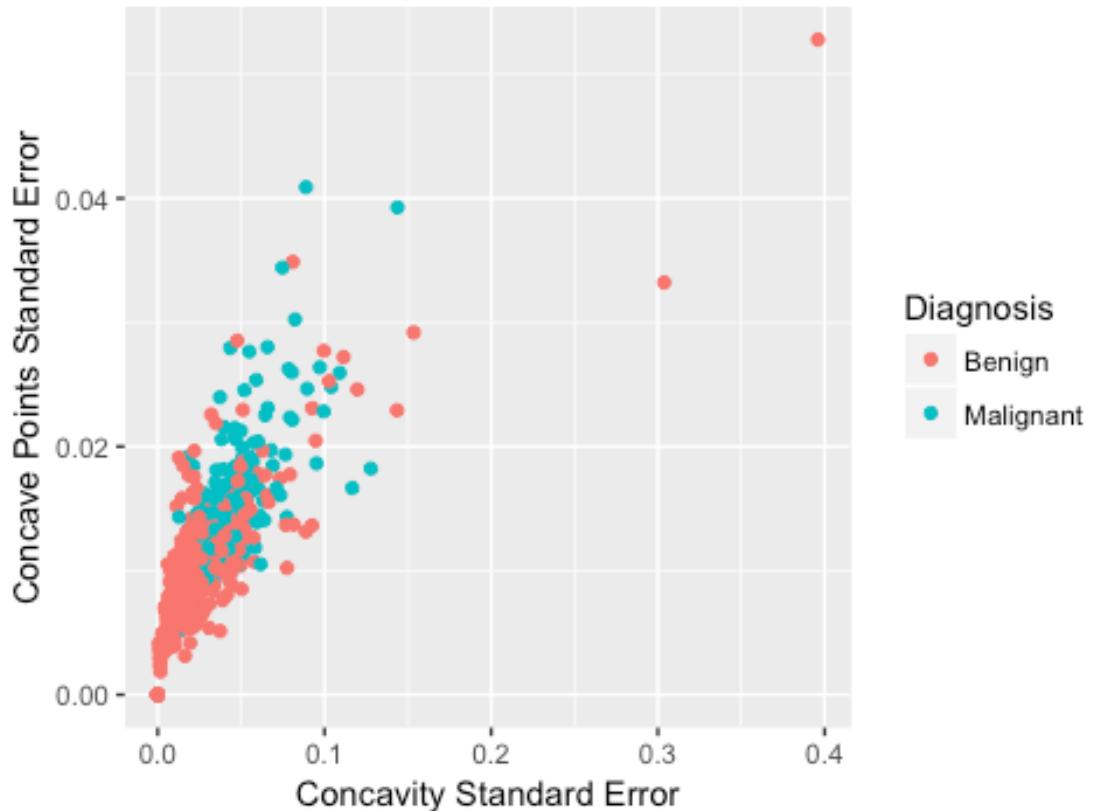
```
#compactness_se vs. concavity_se
concavity_compactness_se_scatter <- ggplot(data = cancer, aes(x = concavity_se, y = compactness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Standard Error", y = "Compactness Standard Error", title = "Compactness Standard Error vs. Concavity Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_compactness_se_scatter
```

Compactness Standard Error vs. Concavity Standard Error



```
#concave.points_se vs. concavity_se
concavity_concave.points_se_scatter <- ggplot(data = cancer, aes(x = concavity_se, y = concave.points_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Standard Error", y = "Concave Points Standard Error", title = "Concave Points Standard Error vs. Concavity Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_concave.points_se_scatter
```

Concave Points Standard Error vs. Concavity Standard Error



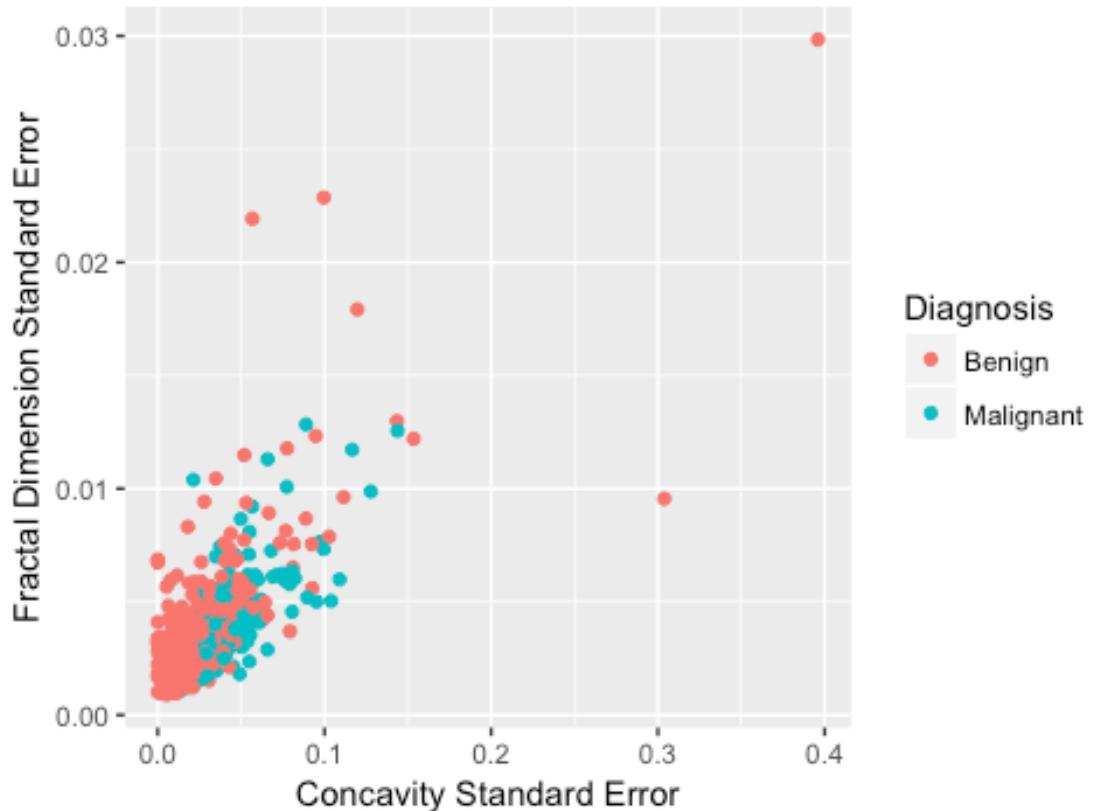
```
#symmetry_se vs. concavity_se
concavity_symmetry_se_scatter <- ggplot(data = cancer, aes(x = concavity_se,
y = symmetry_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Standard Error", y = "Symmetry Standard
Error", title = "Symmetry Standard Error vs. Concavity Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_symmetry_se_scatter
```

Symmetry Standard Error vs. Concavity Standard Error



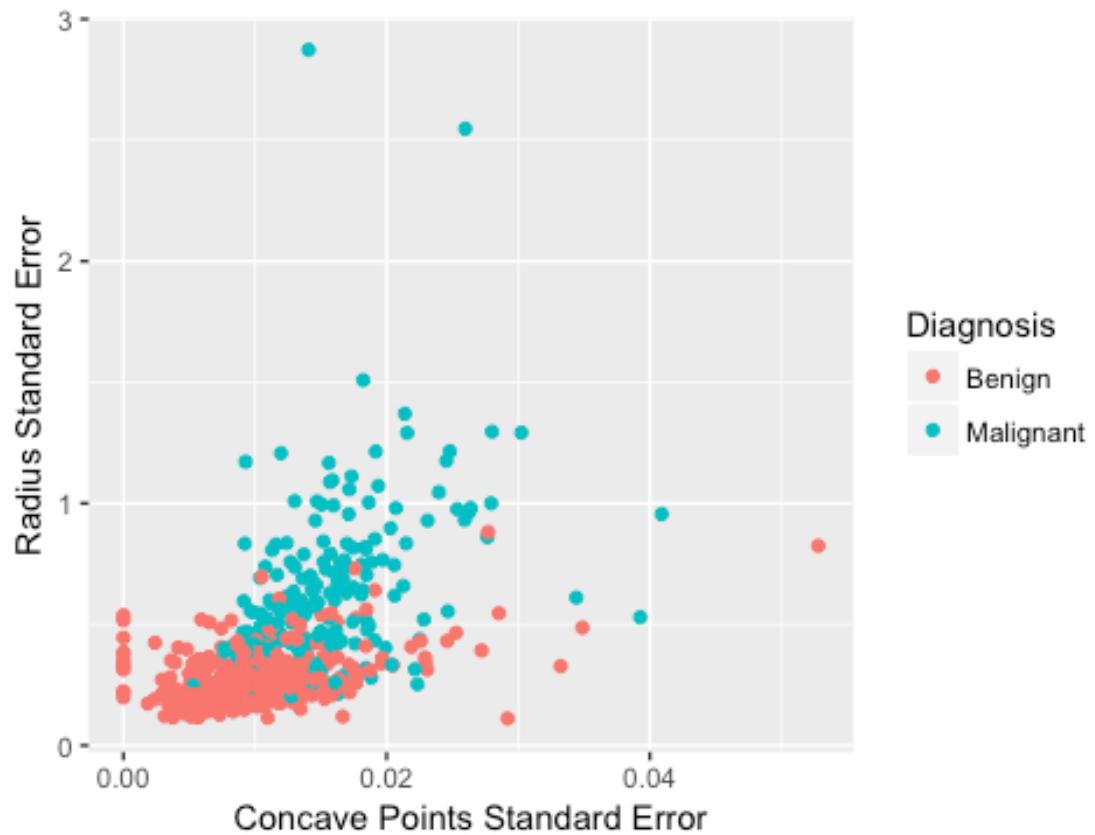
```
#fractal_dimension_se vs. concavity_se
concavity_fractal_dimension_se_scatter <- ggplot(data = cancer, aes(x = concavity_se, y = fractal_dimension_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concavity Standard Error", y = "Fractal Dimension Standard Error", title = "Fractal Dimension Standard Error vs. Concavity Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concavity_fractal_dimension_se_scatter
```

Fractal Dimension Standard Error vs. Concavity Standard Error



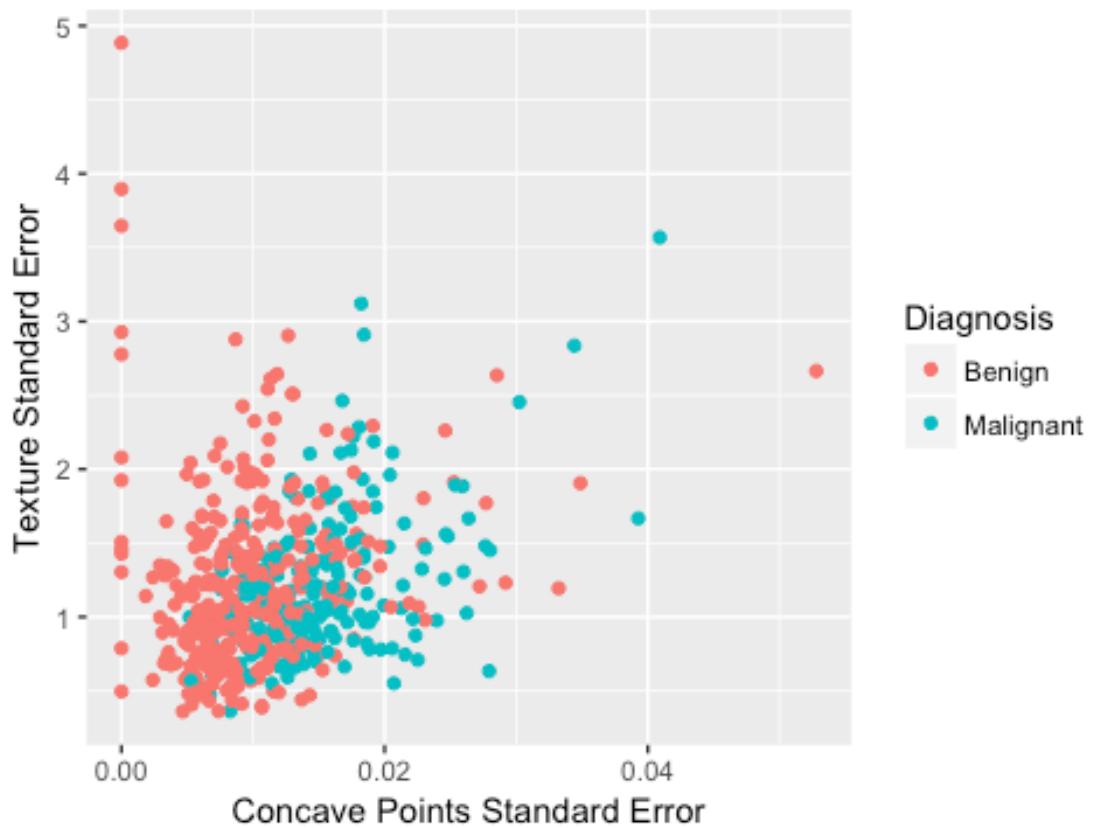
```
#radius_se vs. concave.points_se
concave.points_radius_se_scatter <- ggplot(data = cancer, aes(x = concave.points_se, y = radius_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Standard Error", y = "Radius Standard Error", title = "Radius Standard Error vs. Concave Points Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_radius_se_scatter
```

Radius Standard Error vs. Concave Points Standard Error



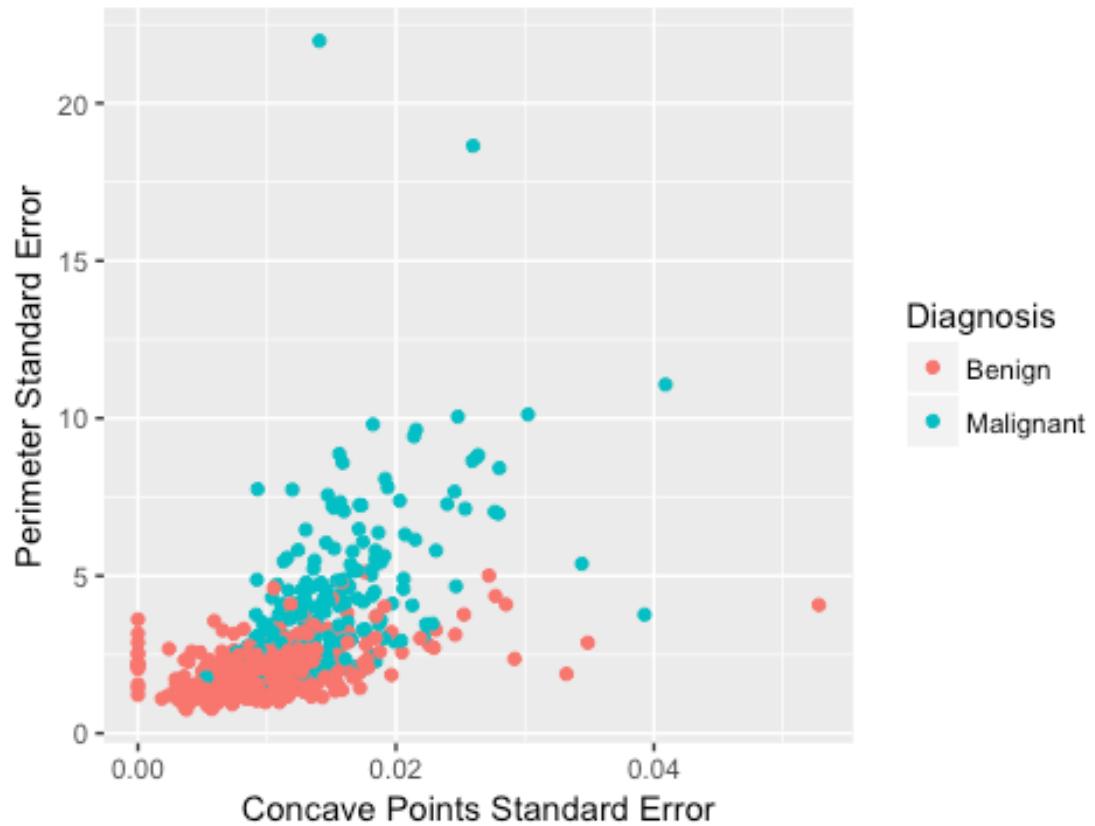
```
#texture_se vs. concave.points_se
concave.points_texture_se_scatter <- ggplot(data = cancer, aes(x = concave.points_se, y = texture_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Standard Error", y = "Texture Standard Error", title = "Texture Standard Error vs. Concave Points Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_texture_se_scatter
```

Texture Standard Error vs. Concave Points Standard Error

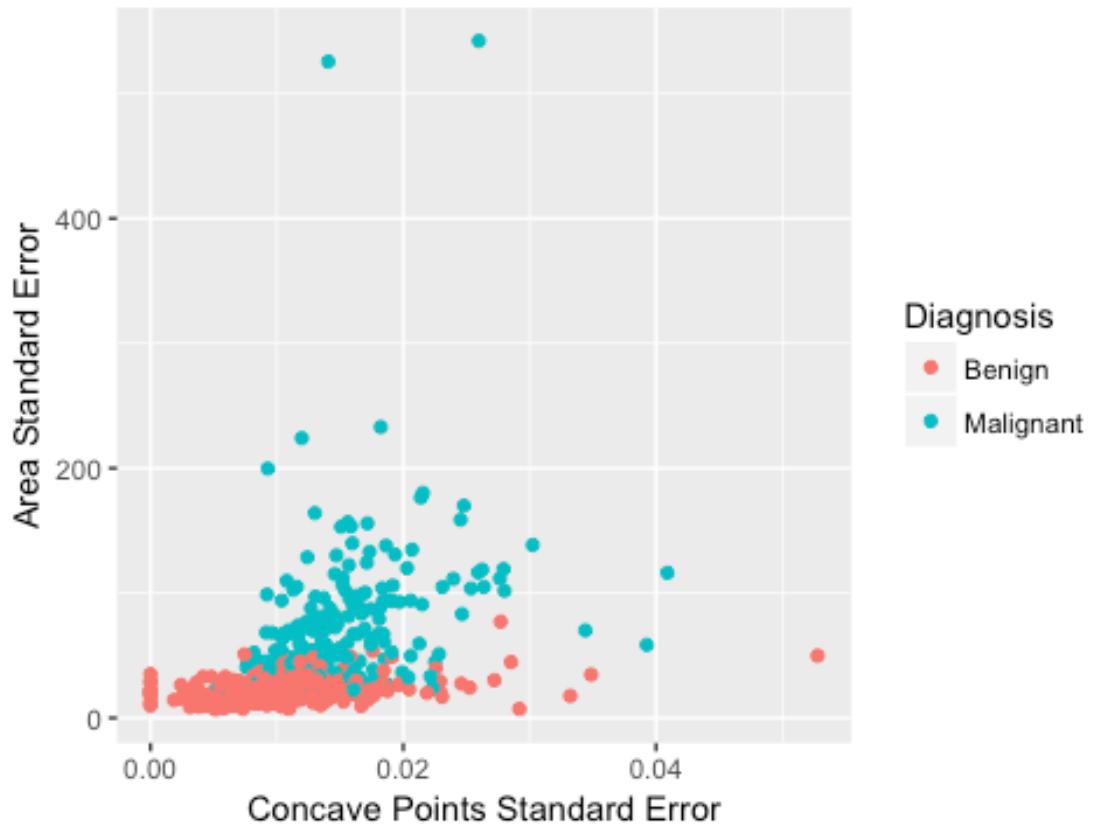


```
#perimeter_se vs. concave.points_se
concave.points_perimeter_se_scatter <- ggplot(data = cancer, aes(x = concave.
points_se, y = perimeter_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Standard Error", y = "Perimeter Standard Error", title = "Perimeter Standard Error vs. Concave Points Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_perimeter_se_scatter
```

Perimeter Standard Error vs. Concave Points Standard Error

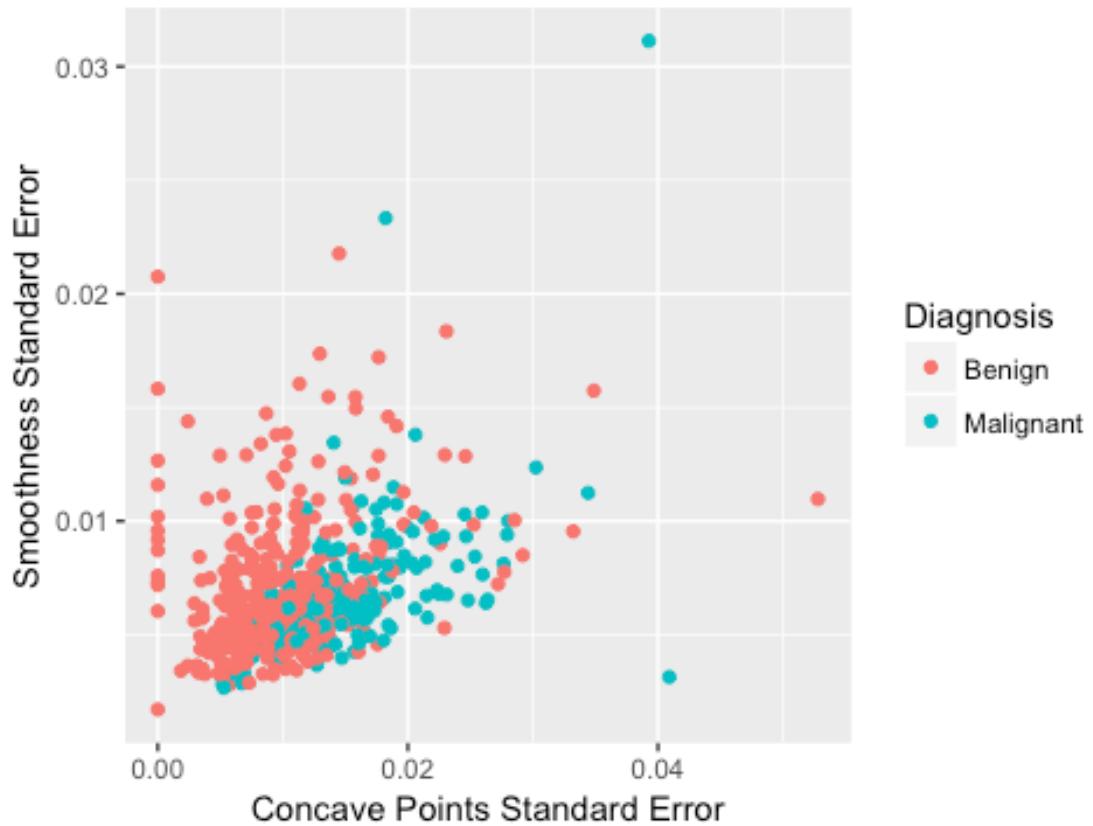


Area Standard Error vs. Concave Points Standard Error



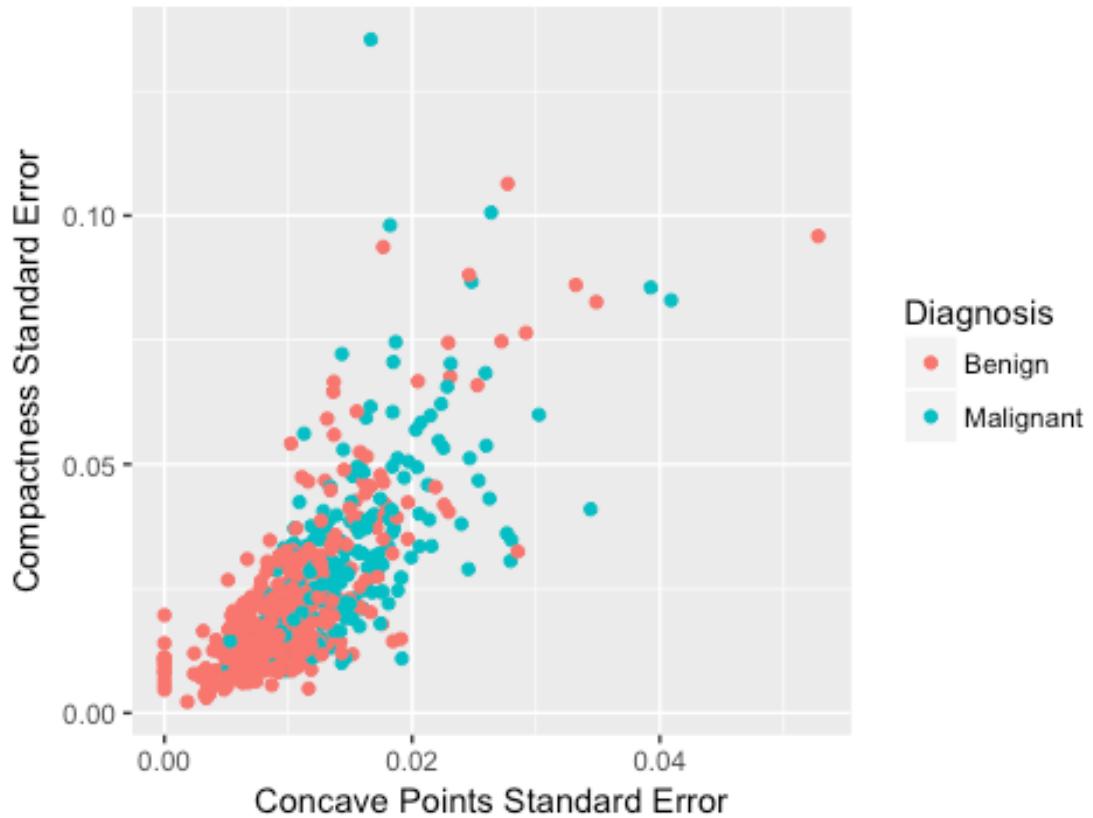
```
#smoothness_se vs. concave.points_se
concave.points_smoothness_se_scatter <- ggplot(data = cancer, aes(x = concave.points_se, y = smoothness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Standard Error", y = "Smoothness Standard Error", title = "Smoothness Standard Error vs. Concave Points Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_smoothness_se_scatter
```

Smoothness Standard Error vs. Concave Points Standard Error

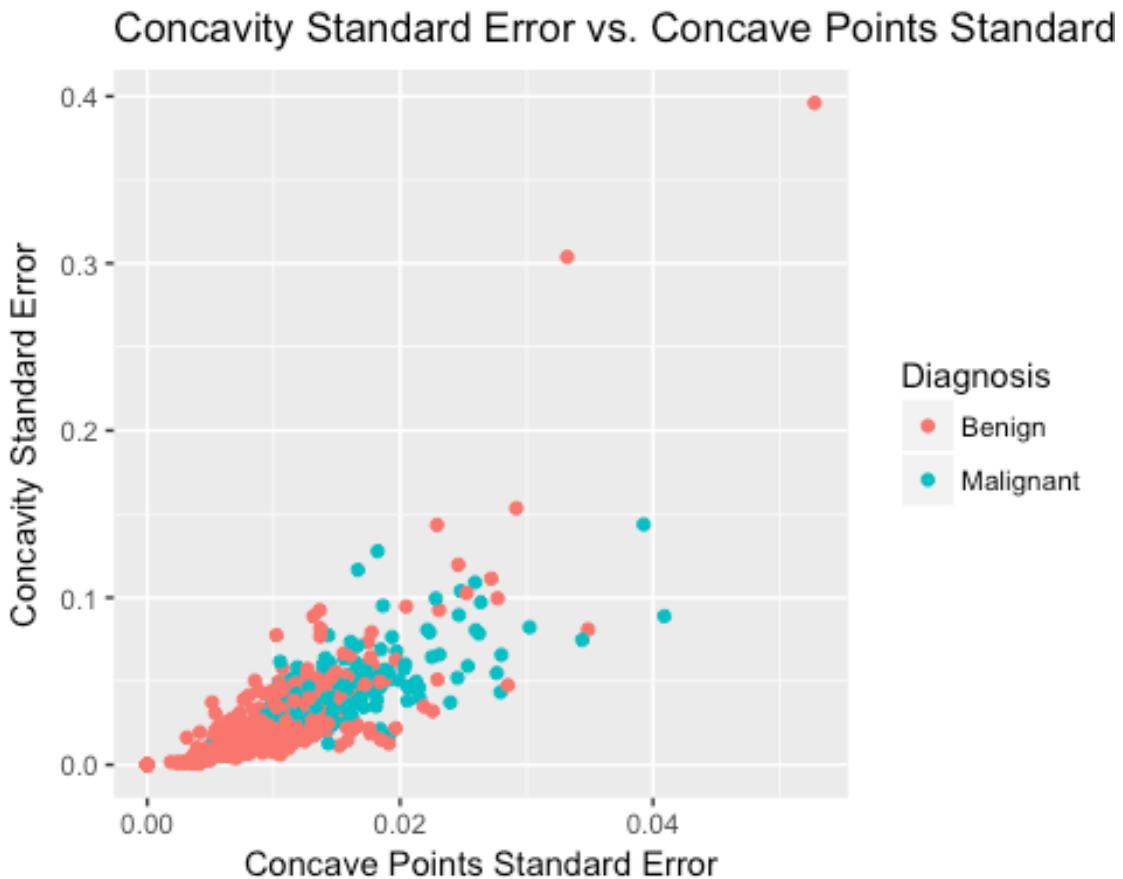


```
#compactness_se vs. concave.points_se
concave.points_compactness_se_scatter <- ggplot(data = cancer, aes(x = concave.points_se, y = compactness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Standard Error", y = "Compactness Standard Error", title = "Compactness Standard Error vs. Concave Points Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_compactness_se_scatter
```

Compactness Standard Error vs. Concave Points Stan

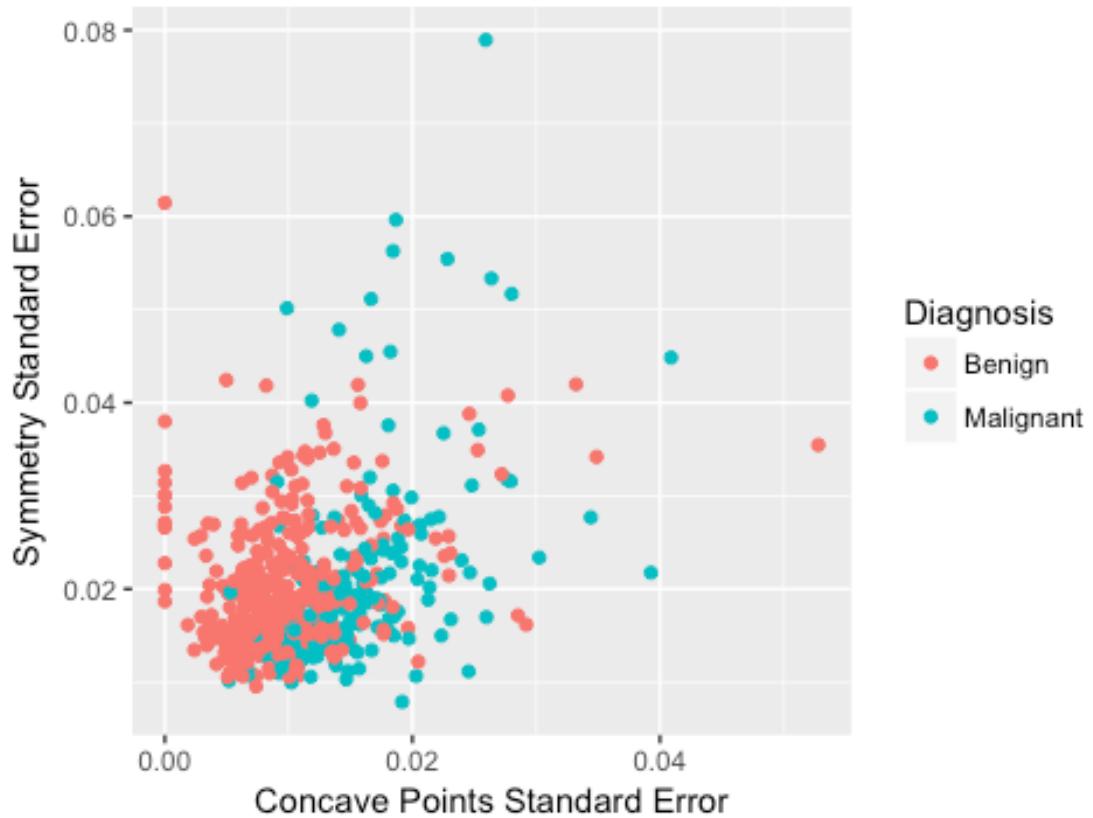


```
#concavity_se vs. concave.points_se
concave.points_concavity_se_scatter <- ggplot(data = cancer, aes(x = concave.
points_se, y = concavity_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Standard Error", y = "Concavity Sta
ndard Error", title = "Concavity Standard Error vs. Concave Points Standard E
rror") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_concavity_se_scatter
```



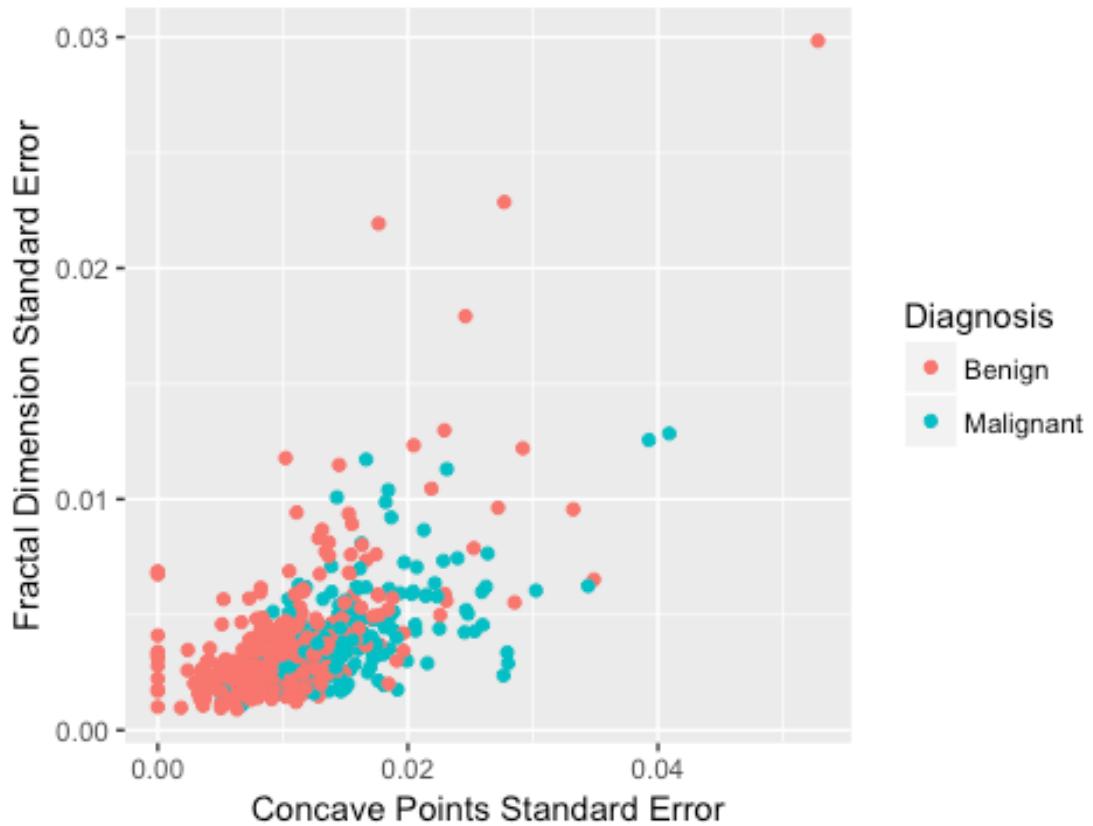
```
#symmetry_se vs. concave.points_se
concave.points_symmetry_se_scatter <- ggplot(data = cancer, aes(x = concave.points_se, y = symmetry_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Standard Error", y = "Symmetry Standard Error", title = "Symmetry Standard Error vs. Concave Points Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_symmetry_se_scatter
```

Symmetry Standard Error vs. Concave Points Standard Error



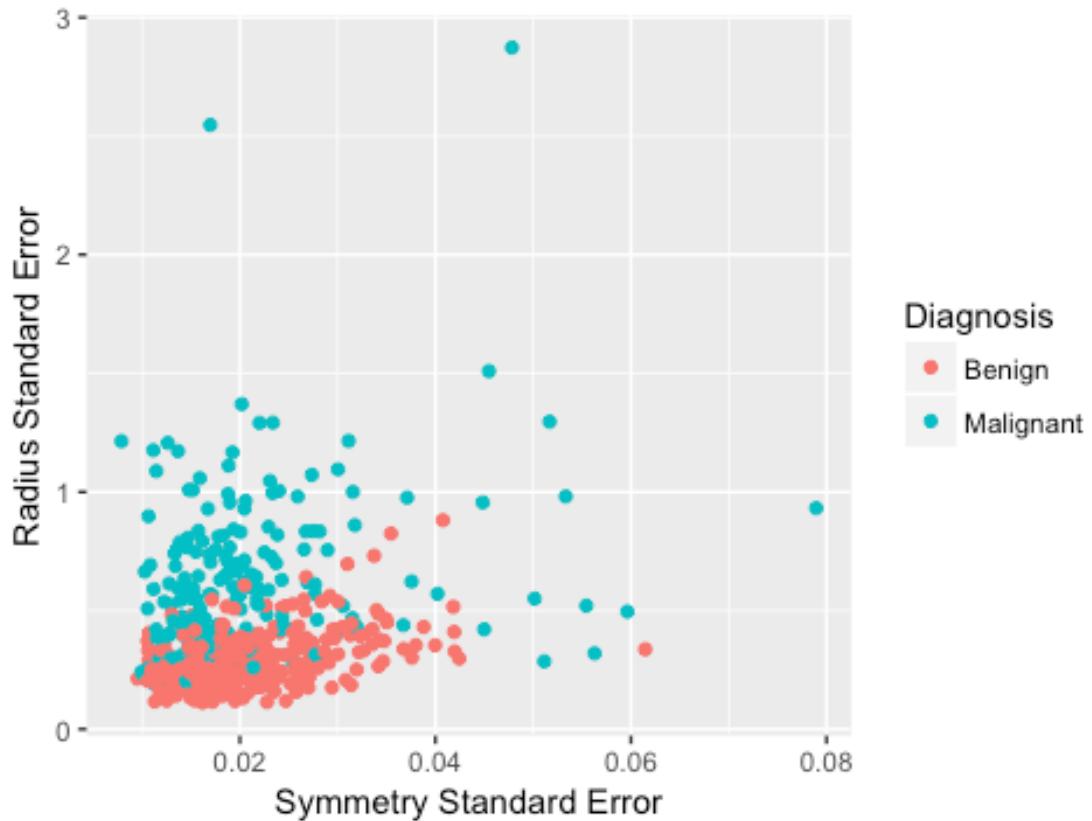
```
#fractal_dimension_se vs. concave.points_se
concave.points_fractal_dimension_se_scatter <- ggplot(data = cancer, aes(x = concave.points_se, y = fractal_dimension_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Concave Points Standard Error", y = "Fractal Dimension Standard Error", title = "Fractal Dimension Standard Error vs. Concave Points Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
concave.points_fractal_dimension_se_scatter
```

Fractal Dimension Standard Error vs. Concave Points :



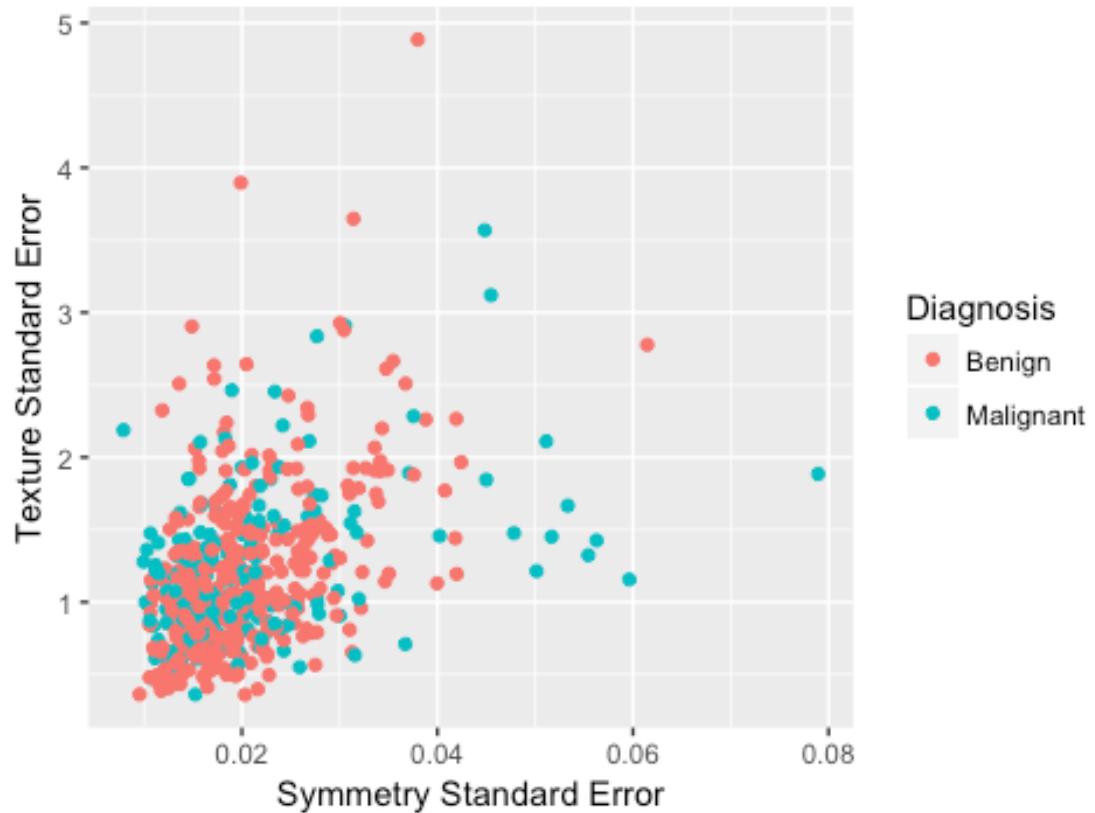
```
#radius_se vs. symmetry_se
symmetry_radius_se_scatter <- ggplot(data = cancer, aes(x = symmetry_se, y = radius_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Standard Error", y = "Radius Standard Error", title = "Radius Standard Error vs. Symmetry Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_radius_se_scatter
```

Radius Standard Error vs. Symmetry Standard Error



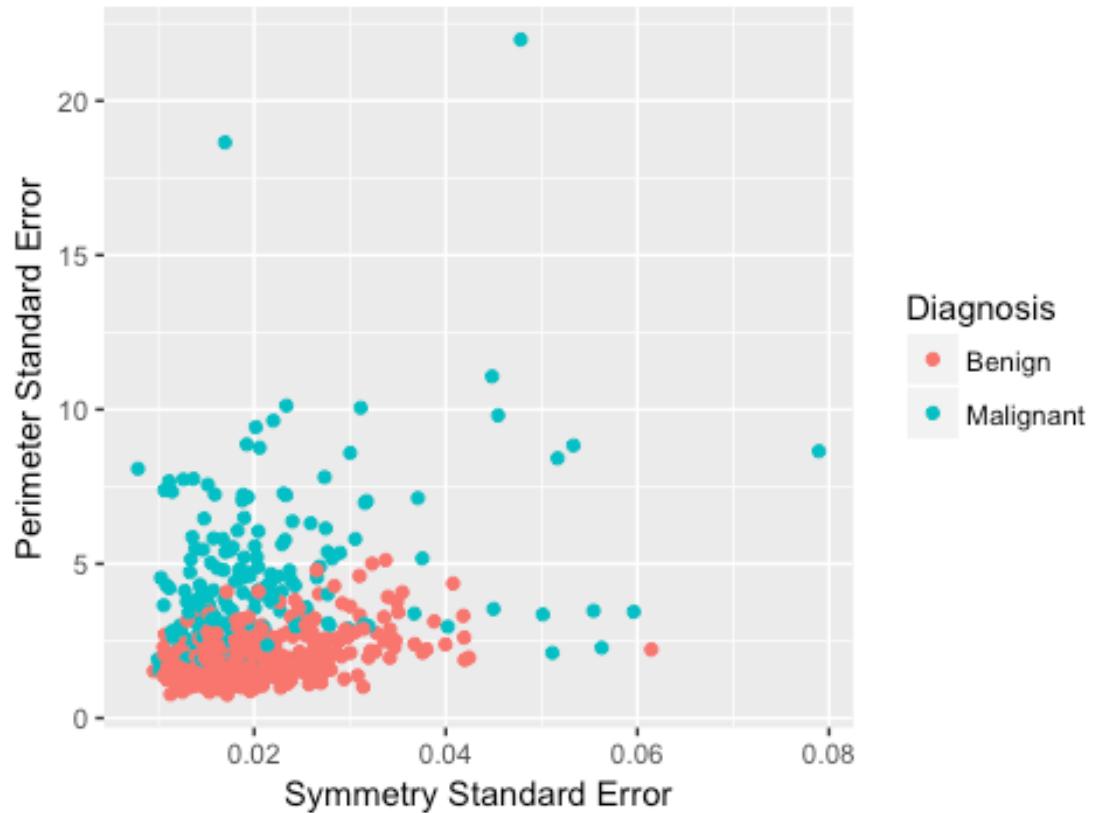
```
#texture_se vs. symmetry_se
symmetry_texture_se_scatter <- ggplot(data = cancer, aes(x = symmetry_se, y = texture_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Standard Error", y = "Texture Standard Error", title = "Texture Standard Error vs. Symmetry Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_texture_se_scatter
```

Texture Standard Error vs. Symmetry Standard Error



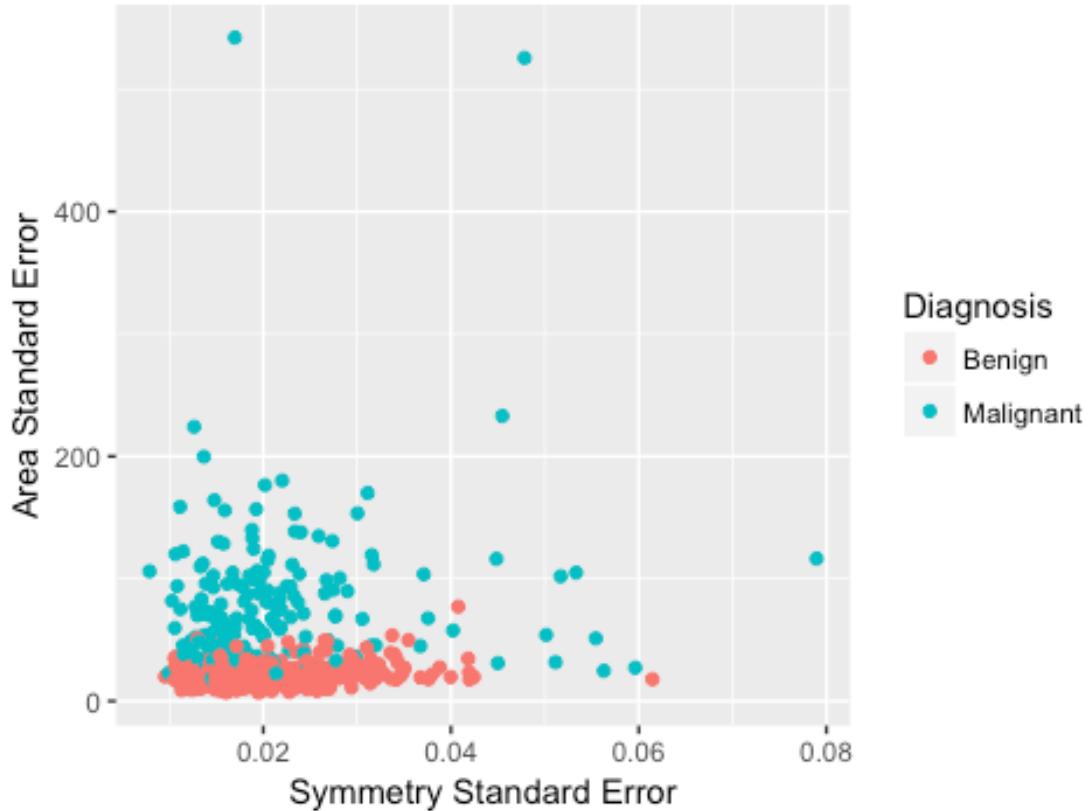
```
#perimeter_se vs. symmetry_se
symmetry_perimeter_se_scatter <- ggplot(data = cancer, aes(x = symmetry_se, y = perimeter_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Standard Error", y = "Perimeter Standard Error", title = "Perimeter Standard Error vs. Symmetry Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_perimeter_se_scatter
```

Perimeter Standard Error vs. Symmetry Standard Error



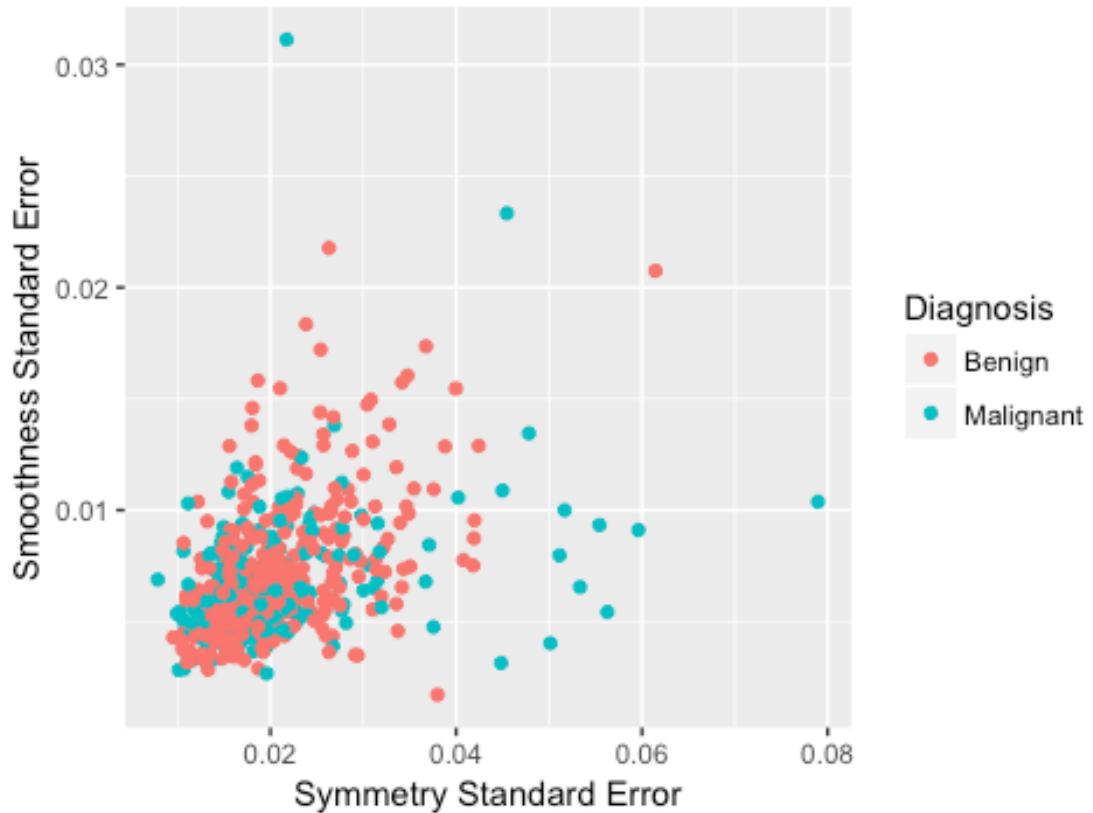
```
#area_se vs. symmetry_se
symmetry_area_se_scatter <- ggplot(data = cancer, aes(x = symmetry_se, y = area_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Standard Error", y = "Area Standard Error",
  , title = "Area Standard Error vs. Symmetry Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_area_se_scatter
```

Area Standard Error vs. Symmetry Standard Error



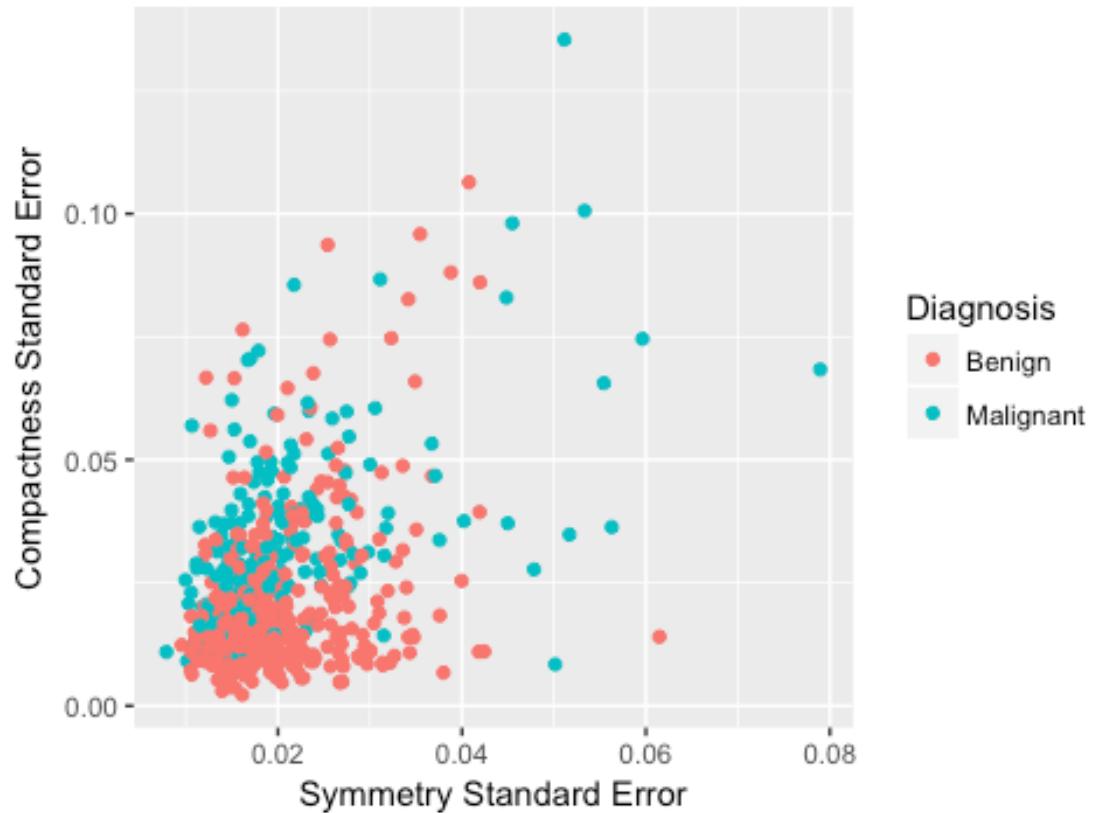
```
#smoothness_se vs. symmetry_se
symmetry_smoothness_se_scatter <- ggplot(data = cancer, aes(x = symmetry_se,
y = smoothness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Standard Error", y = "Smoothness Standard
Error", title = "Smoothness Standard Error vs. Symmetry Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_smoothness_se_scatter
```

Smoothness Standard Error vs. Symmetry Standard Error



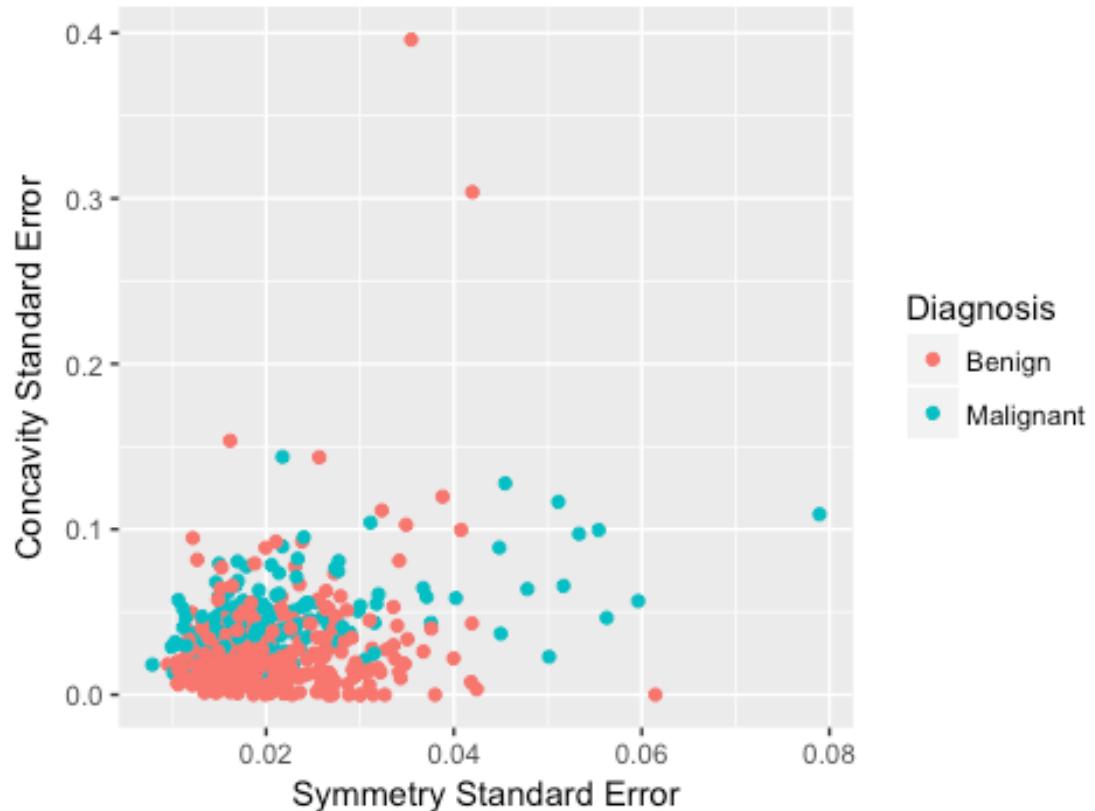
```
#compactness_se vs. symmetry_se
symmetry_compactness_se_scatter <- ggplot(data = cancer, aes(x = symmetry_se,
y = compactness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Standard Error", y = "Compactness Standard Error", title = "Compactness Standard Error vs. Symmetry Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_compactness_se_scatter
```

Compactness Standard Error vs. Symmetry Standard Error



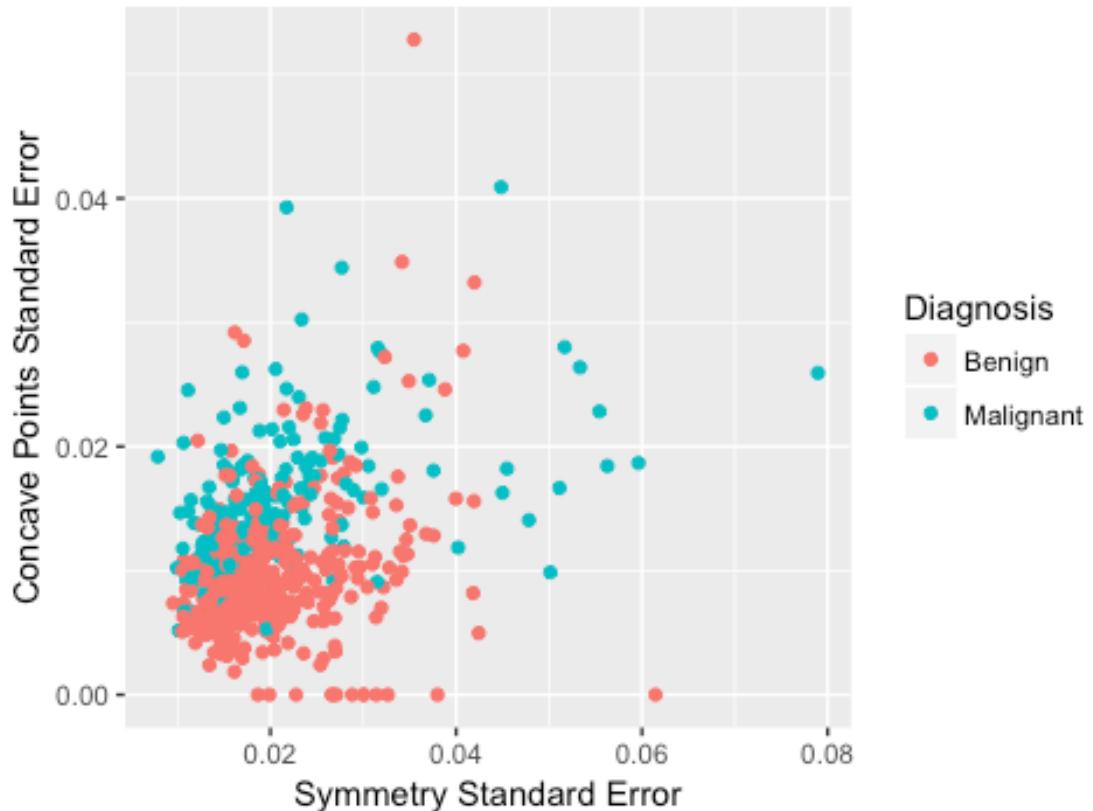
```
#concavity_se vs. symmetry_se
symmetry_concavity_se_scatter <- ggplot(data = cancer, aes(x = symmetry_se, y = concavity_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Standard Error", y = "Concavity Standard Error", title = "Concavity Standard Error vs. Symmetry Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_concavity_se_scatter
```

Concavity Standard Error vs. Symmetry Standard Error



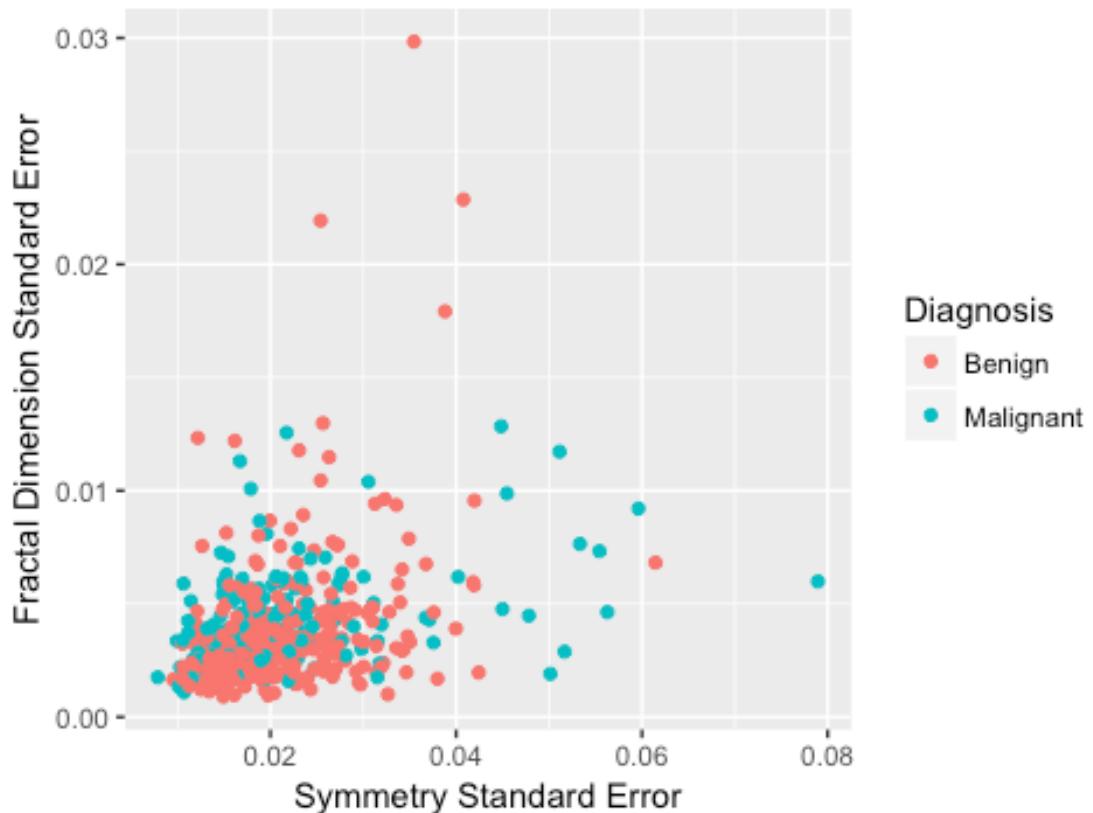
```
#concave.points_se vs. symmetry_se
symmetry_concave.points_se_scatter <- ggplot(data = cancer, aes(x = symmetry_se, y = concave.points_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Standard Error", y = "Concave Points Standard Error", title = "Concave Points Standard Error vs. Symmetry Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_concave.points_se_scatter
```

Concave Points Standard Error vs. Symmetry Standard Error



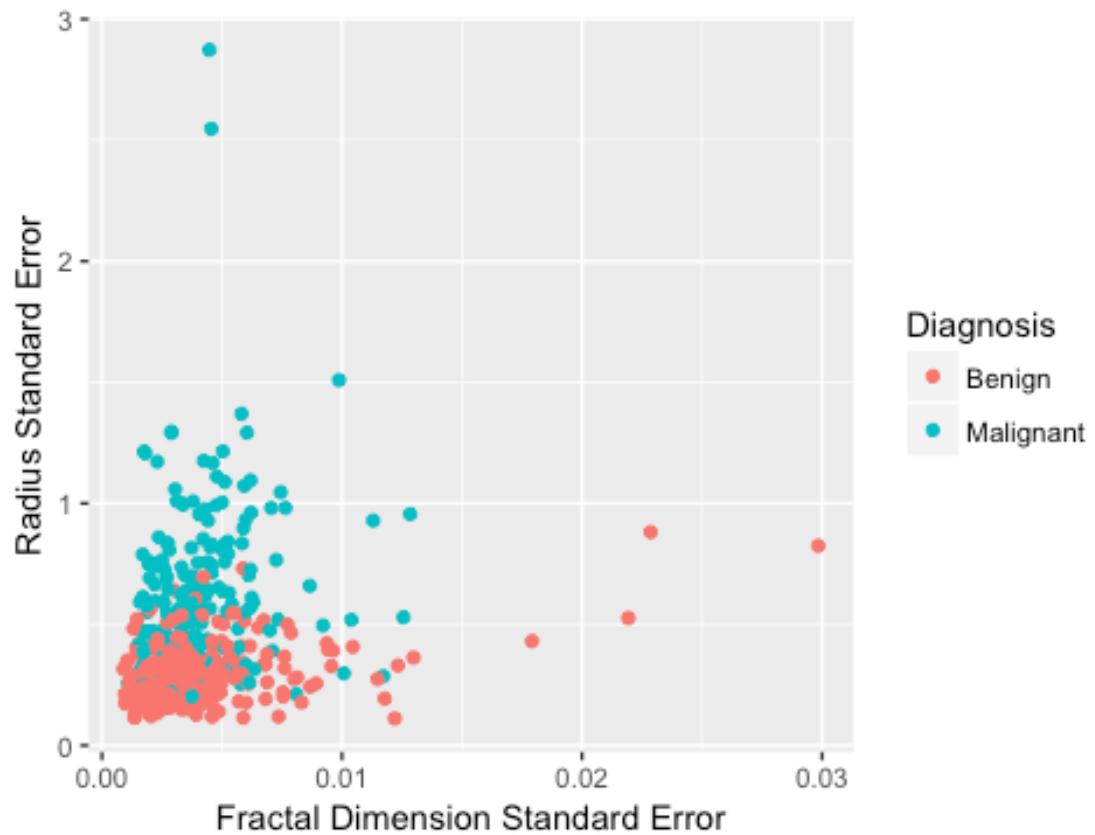
```
#fractal_dimension_se vs. symmetry_se
symmetry_fractal_dimension_se_scatter <- ggplot(data = cancer, aes(x = symmetry_se, y = fractal_dimension_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Symmetry Standard Error", y = "Fractal Dimension Standard Error", title = "Fractal Dimension Standard Error vs. Symmetry Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
symmetry_fractal_dimension_se_scatter
```

Fractal Dimension Standard Error vs. Symmetry Standard Error



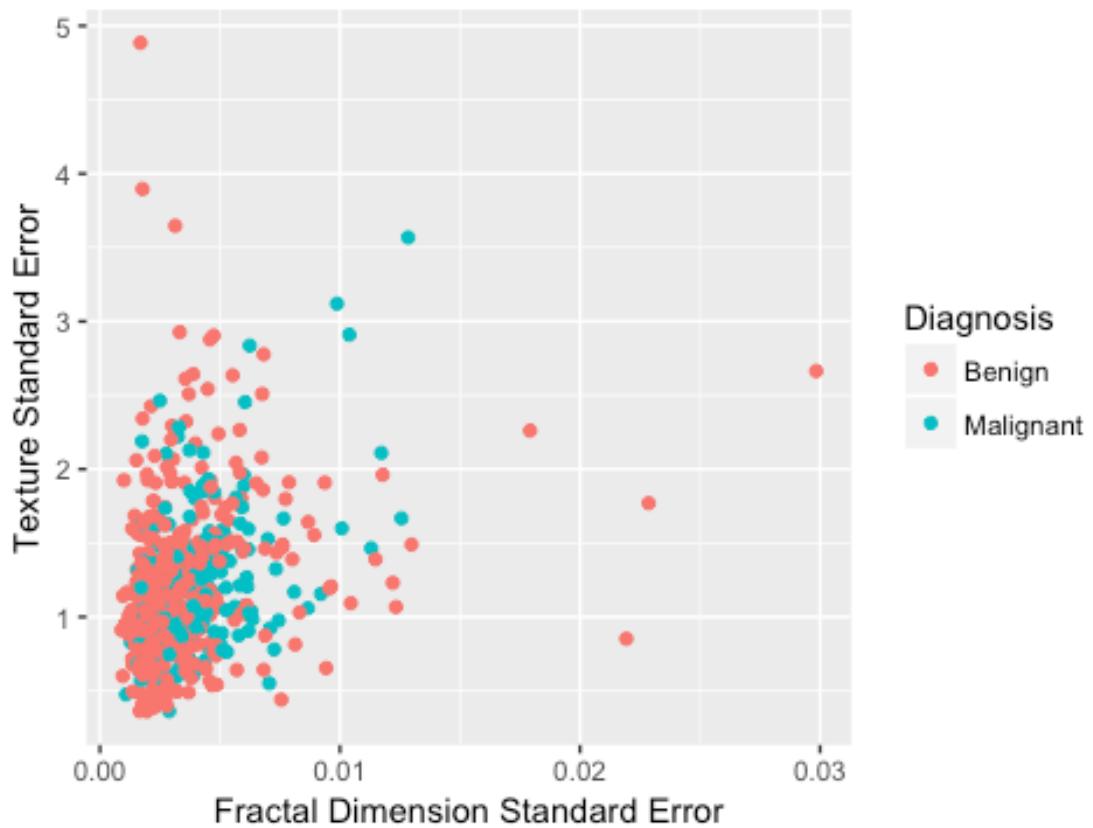
```
#radius_se vs. fractal_dimension_se
fractal_dimension_radius_se_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_se, y = radius_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Standard Error", y = "Radius Standard Error", title = "Radius Standard Error vs. Fractal Dimension Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_radius_se_scatter
```

Radius Standard Error vs. Fractal Dimension Standard E



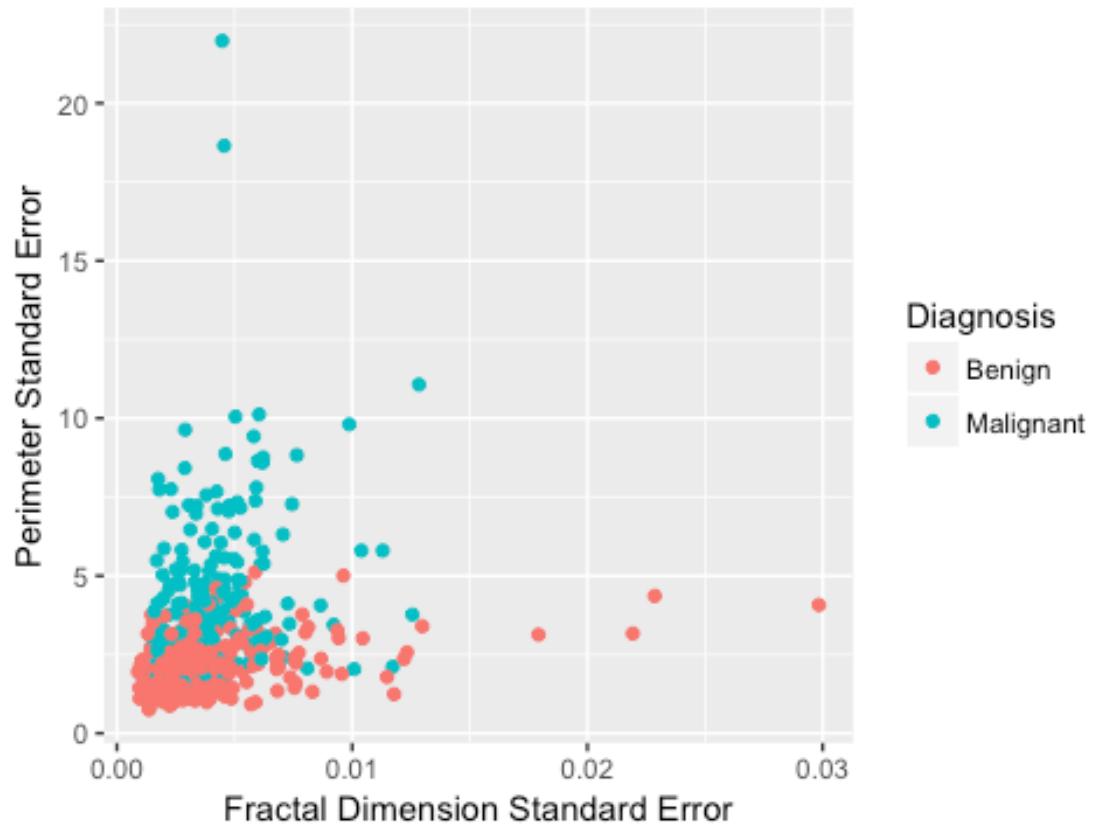
```
#texture_se vs. fractal_dimension_se
fractal_dimension_texture_se_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_se, y = texture_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Standard Error", y = "Texture Standard Error", title = "Texture Standard Error vs. Fractal Dimension Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_texture_se_scatter
```

Texture Standard Error vs. Fractal Dimension Standard E



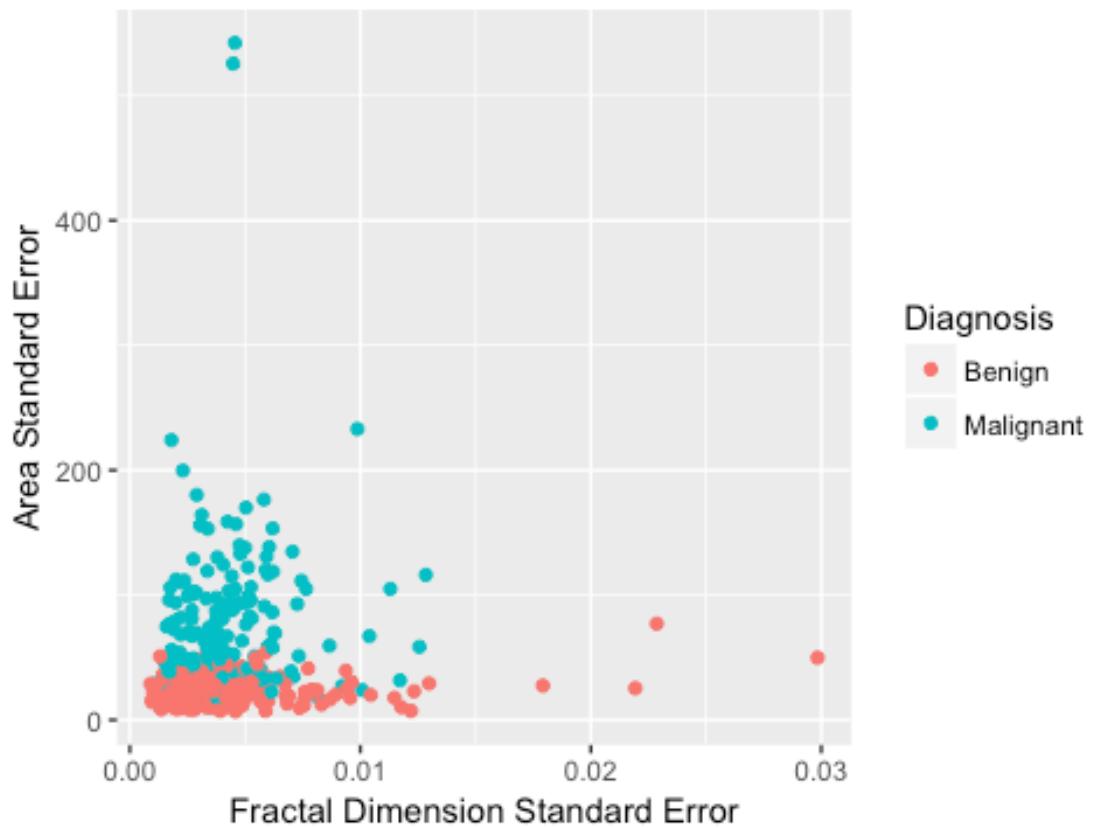
```
#perimeter_se vs. fractal_dimension_se
fractal_dimension_perimeter_se_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_se, y = perimeter_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Standard Error", y = "Perimeter Standard Error", title = "Perimeter Standard Error vs. Fractal Dimension Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_perimeter_se_scatter
```

Perimeter Standard Error vs. Fractal Dimension Standard Error



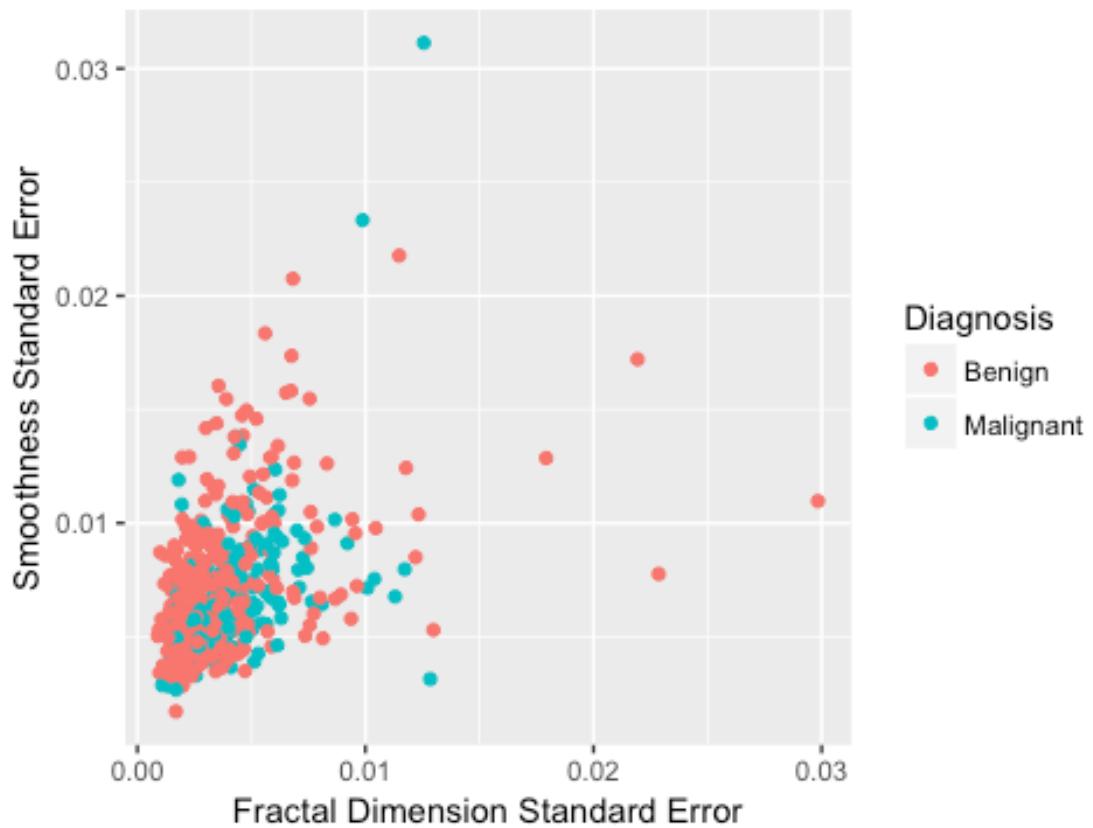
```
#area_se vs. fractal_dimension_se
fractal_dimension_area_se_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_se, y = area_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Standard Error", y = "Area Standard Error", title = "Area Standard Error vs. Fractal Dimension Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_area_se_scatter
```

Area Standard Error vs. Fractal Dimension Standard Er



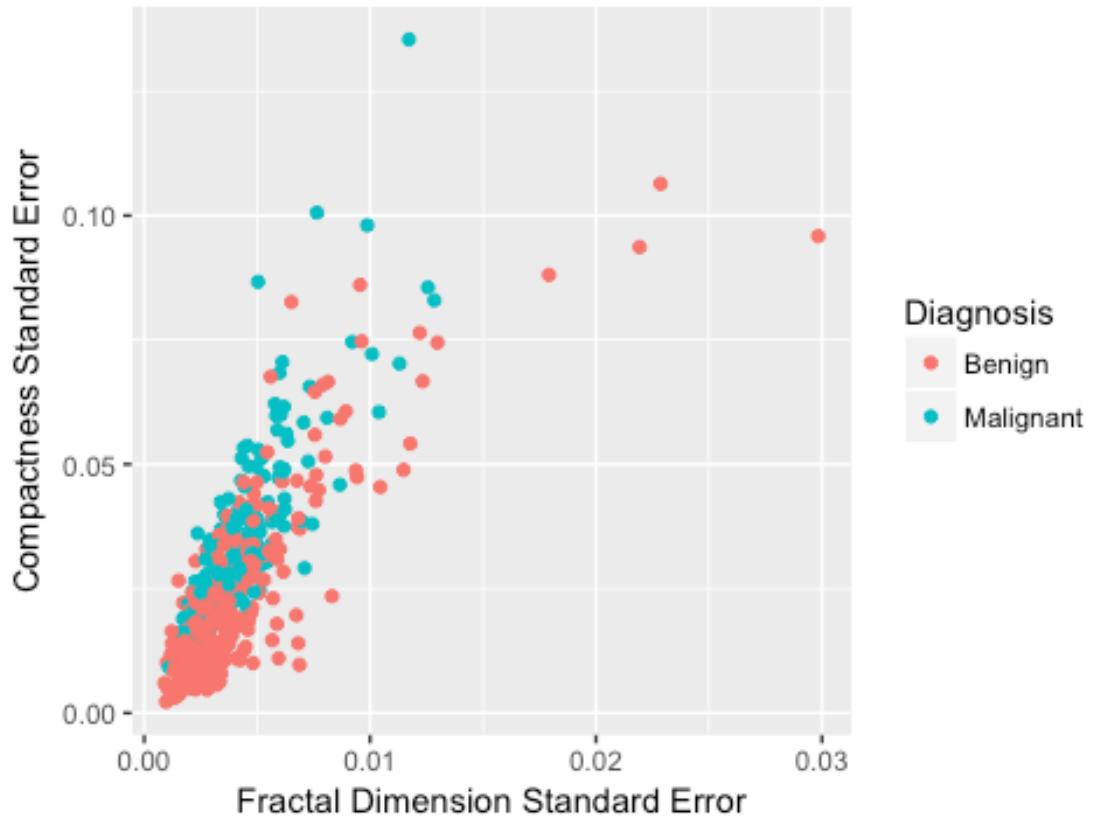
```
#smoothness_se vs. fractal_dimension_se
fractal_dimension_smoothness_se_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_se, y = smoothness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Standard Error", y = "Smoothness Standard Error", title = "Smoothness Standard Error vs. Fractal Dimension Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_smoothness_se_scatter
```

Smoothness Standard Error vs. Fractal Dimension Standard Error



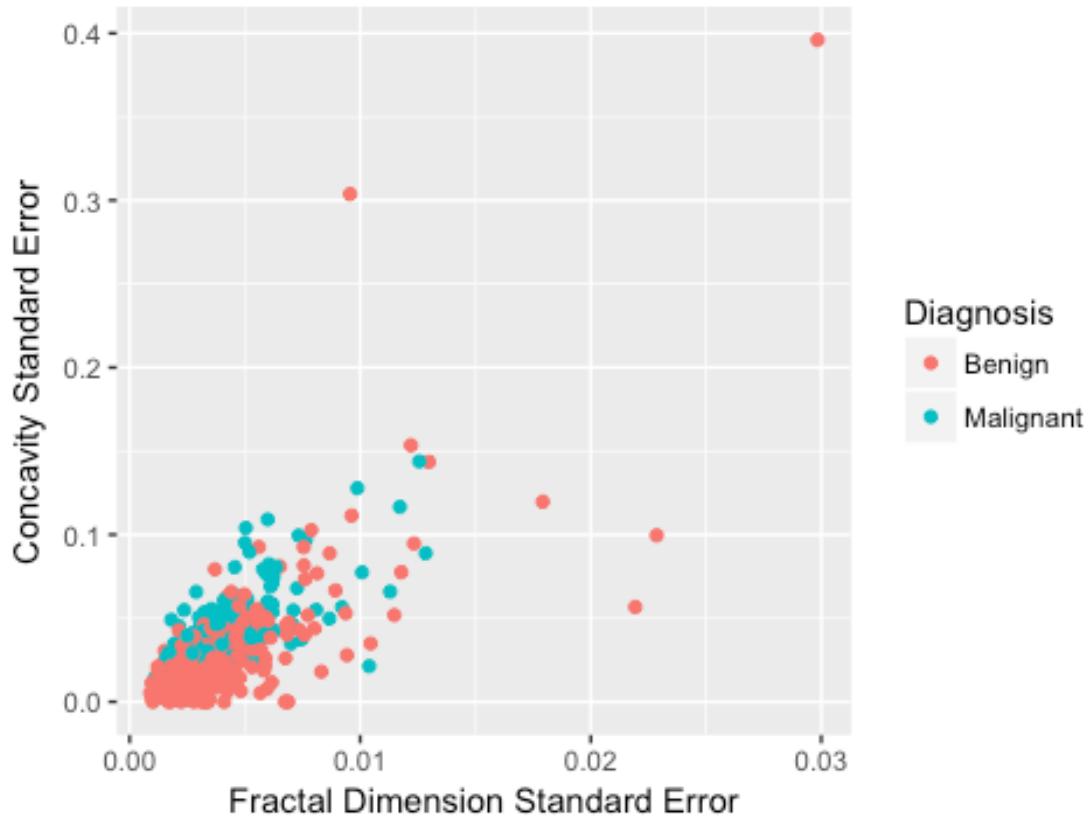
```
#compactness_se vs. fractal_dimension_se
fractal_dimension_compactness_se_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_se, y = compactness_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Standard Error", y = "Compactness Standard Error", title = "Compactness Standard Error vs. Fractal Dimension Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_compactness_se_scatter
```

Compactness Standard Error vs. Fractal Dimension Standard Error



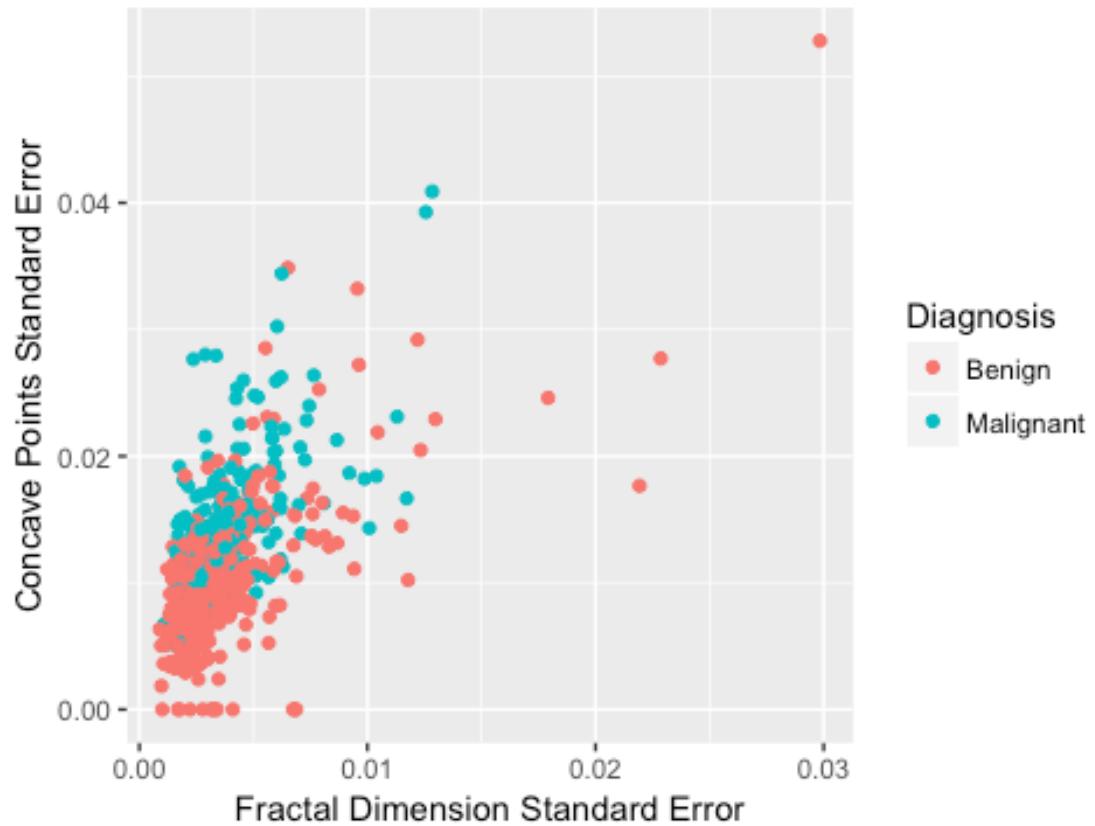
```
#concavity_se vs. fractal_dimension_se
fractal_dimension_concavity_se_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_se, y = concavity_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Standard Error", y = "Concavity Standard Error", title = "Concavity Standard Error vs. Fractal Dimension Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_concavity_se_scatter
```

Concavity Standard Error vs. Fractal Dimension Standard Error

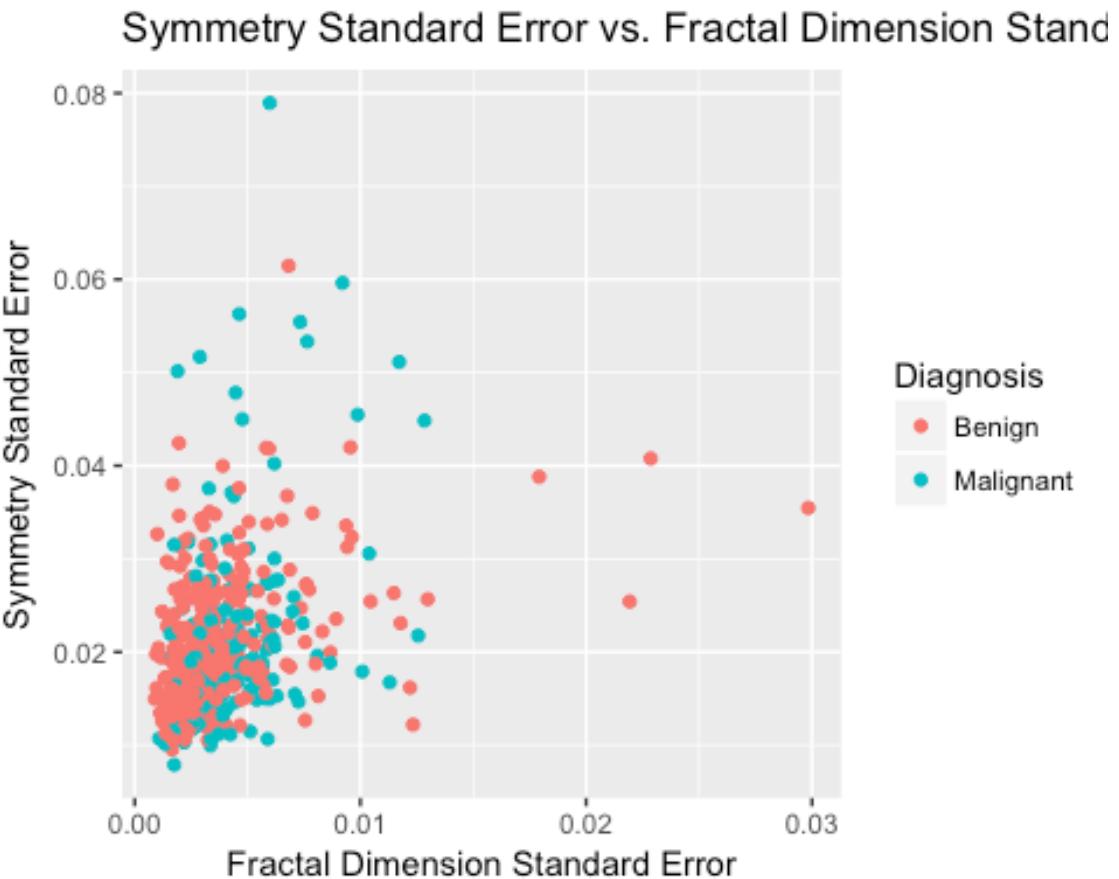


```
#concave.points_se vs. fractal_dimension_se
fractal_dimension_concave.points_se_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_se, y = concave.points_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Standard Error", y = "Concave Points Standard Error", title = "Concave Points Standard Error vs. Fractal Dimension Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_concave.points_se_scatter
```

Concave Points Standard Error vs. Fractal Dimension :



```
#symmetry_se vs. fractal_dimension_se
fractal_dimension_symmetry_se_scatter <- ggplot(data = cancer, aes(x = fractal_dimension_se, y = symmetry_se, color = as.factor(diagnosis))) +
  geom_point() + labs(x = "Fractal Dimension Standard Error", y = "Symmetry Standard Error", title = "Symmetry Standard Error vs. Fractal Dimension Standard Error") +
  scale_color_discrete(name = "Diagnosis", labels = c("Benign", "Malignant"))
fractal_dimension_symmetry_se_scatter
```



#3 - Split the data into test and training data.

```

# create new dataframe by copying cancer data
cancer_tree_all <- cancer
# make the diagnosis variable binary, malignant = 1, benign = 0 (factors)
cancer_tree_all$diagnosis <- ifelse(cancer_tree_all$diagnosis=='M', 1, 0)
cancer_tree_all$diagnosis <- as.factor(cancer_tree_all$diagnosis)

# set the seed
set.seed(1842)
# save the number of rows in data as n
n <- nrow(cancer_tree_all)
# create the test_data dataset of 20% of the data number of rows
test_index <- sample.int(n, size = round(0.2*n))
test_data_tree <- cancer_tree_all[test_index,]
# create the train_data dataset as the remaining 80%
# number of rows
train_data_tree <- cancer_tree_all[-test_index, ]
# glimpse of both datasets
glimpse(test_data_tree)

## Observations: 114
## Variables: 33

```

```

## $ id                               <int> 8712766, 91903901, 8811842, 89143602, ...
## $ diagnosis                         <fct> 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, ...
## $ radius_mean                        <dbl> 17.470, 11.670, 19.800, 14.410, 18.450...
## $ texture_mean                       <dbl> 24.68, 20.02, 21.56, 19.73, 21.91, 23....
## $ perimeter_mean                     <dbl> 116.10, 75.21, 129.70, 96.03, 120.20, ...
## $ area_mean                           <dbl> 984.6, 416.2, 1230.0, 651.0, 1075.0, 8...
## $ smoothness_mean                     <dbl> 0.10490, 0.10160, 0.09383, 0.08757, 0....
## $ compactness_mean                   <dbl> 0.16030, 0.09453, 0.13060, 0.16760, 0....
## $ concavity_mean                     <dbl> 0.21590, 0.04200, 0.12720, 0.13620, 0....
## $ concave.points_mean               <dbl> 0.10430, 0.02157, 0.08691, 0.06602, 0....
## $ symmetry_mean                      <dbl> 0.1538, 0.1859, 0.2094, 0.1714, 0.1692...
## $ fractal_dimension_mean             <dbl> 0.06365, 0.06461, 0.05581, 0.07192, 0....
## $ radius_se                           <dbl> 1.0880, 0.2067, 0.9553, 0.8811, 0.5959...
## $ texture_se                          <dbl> 1.4100, 0.8745, 1.1860, 1.7700, 1.2020...
## $ perimeter_se                        <dbl> 7.337, 1.393, 6.487, 4.360, 3.766, 2.9...
## $ area_se                            <dbl> 122.30, 15.34, 124.40, 77.11, 68.35, 4...
## $ smoothness_se                       <dbl> 0.006174, 0.005251, 0.006804, 0.007762...
## $ compactness_se                      <dbl> 0.036340, 0.017270, 0.031690, 0.106400...
## $ concavity_se                        <dbl> 0.046440, 0.018400, 0.034460, 0.099600...
## $ concave.points_se                  <dbl> 0.015690, 0.005298, 0.017120, 0.027710...
## $ symmetry_se                         <dbl> 0.011450, 0.014490, 0.018970, 0.040770...
## $ fractal_dimension_se                <dbl> 0.005120, 0.002671, 0.004045, 0.022860...
## $ radius_worst                        <dbl> 23.14, 13.35, 25.73, 15.77, 22.52, 19....
## $ texture_worst                       <dbl> 32.33, 28.81, 28.64, 22.13, 31.39, 31....
## $ perimeter_worst                     <dbl> 155.30, 87.00, 170.30, 101.70, 145.60, ...
## $ area_worst                           <dbl> 1660.0, 550.6, 2009.0, 767.3, 1590.0, ...
## $ smoothness_worst                    <dbl> 0.13760, 0.15500, 0.13530, 0.09983, 0....
## $ compactness_worst                   <dbl> 0.38300, 0.29640, 0.32350, 0.24720, 0....
## $ concavity_worst                     <dbl> 0.48900, 0.27580, 0.36170, 0.22200, 0....
## $ concave.points_worst               <dbl> 0.17210, 0.08120, 0.18200, 0.10210, 0....
## $ symmetry_worst                      <dbl> 0.2160, 0.3206, 0.3070, 0.2272, 0.3109...
## $ fractal_dimension_worst             <dbl> 0.09300, 0.08950, 0.08255, 0.08799, 0....
## $ X                                 <lgcl> NA, NA, NA, NA, NA, NA, NA, NA, NA, ...

glimpse(train_data_tree)

## Observations: 455
## Variables: 33
## $ id                               <int> 842302, 84300903, 84348301, 843786, 84...
## $ diagnosis                         <fct> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ radius_mean                        <dbl> 17.990, 19.690, 11.420, 12.450, 13.710...
## $ texture_mean                       <dbl> 10.38, 21.25, 20.38, 15.70, 20.83, 21....
## $ perimeter_mean                     <dbl> 122.80, 130.00, 77.58, 82.57, 90.20, 8...
## $ area_mean                           <dbl> 1001.0, 1203.0, 386.1, 477.1, 577.9, 5...
## $ smoothness_mean                     <dbl> 0.11840, 0.10960, 0.14250, 0.12780, 0....
## $ compactness_mean                   <dbl> 0.27760, 0.15990, 0.28390, 0.17000, 0....
## $ concavity_mean                     <dbl> 0.30010, 0.19740, 0.24140, 0.15780, 0....
## $ concave.points_mean               <dbl> 0.14710, 0.12790, 0.10520, 0.08089, 0....
## $ symmetry_mean                      <dbl> 0.2419, 0.2069, 0.2597, 0.2087, 0.2196...
## $ fractal_dimension_mean             <dbl> 0.07871, 0.05999, 0.09744, 0.07613, 0....

```

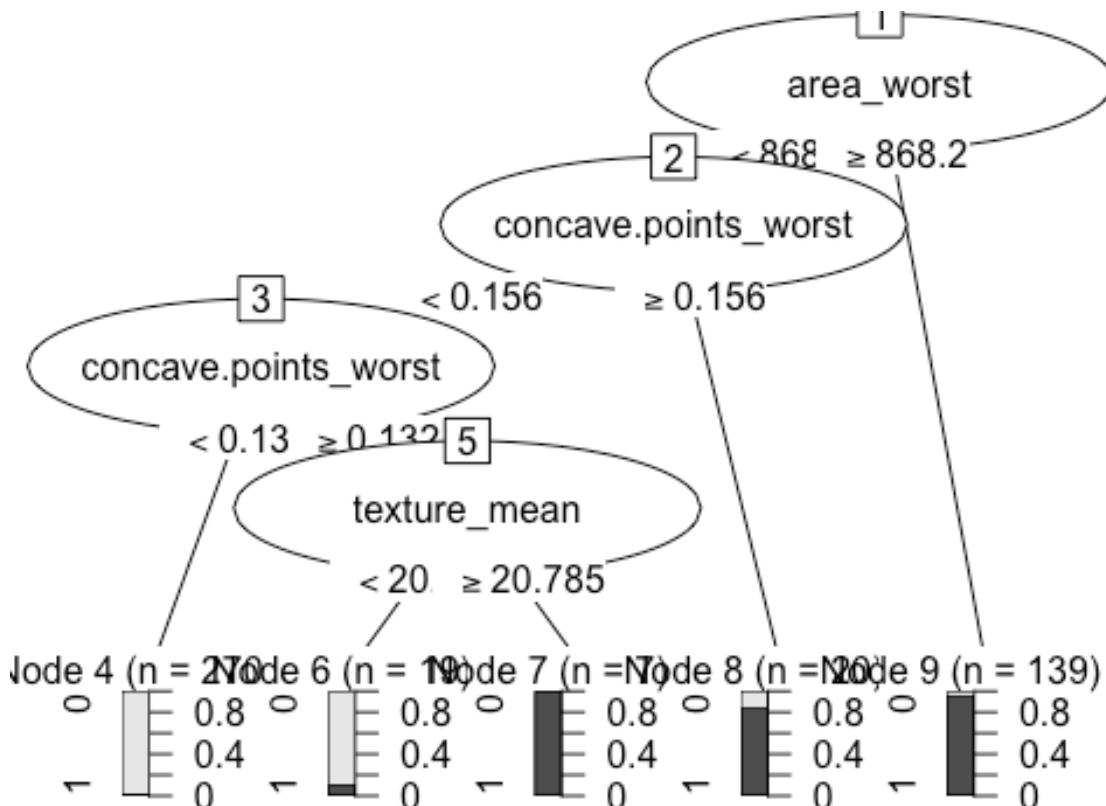


```

## 1) root 455 163 0 (0.64175824 0.35824176)
## 2) area_worst< 868.2 316 30 0 (0.90506329 0.09493671)
## 4) concave.points_worst< 0.1563 296 13 0 (0.95608108 0.04391892)
## 8) concave.points_worst< 0.13235 270 4 0 (0.98518519 0.01481481)
*
## 9) concave.points_worst>=0.13235 26 9 0 (0.65384615 0.34615385)
## 18) texture_mean< 20.785 19 2 0 (0.89473684 0.10526316) *
## 19) texture_mean>=20.785 7 0 1 (0.00000000 1.00000000) *
## 5) concave.points_worst>=0.1563 20 3 1 (0.15000000 0.85000000) *
## 3) area_worst>=868.2 139 6 1 (0.04316547 0.95683453) *

#partykit based plot of the tree produced
plot(as.party(tree_1))

```



```

#determine complexity parameter for tree
printcp(tree_1)

##
## Classification tree:
## rpart(formula = diagnosis ~ radius_mean + texture_mean + perimeter_mean +
##        area_mean + smoothness_mean + compactness_mean + concavity_mean +
##        concave.points_mean + symmetry_mean + fractal_dimension_mean +
##        radius_se + texture_se + perimeter_se + area_se + smoothness_se +

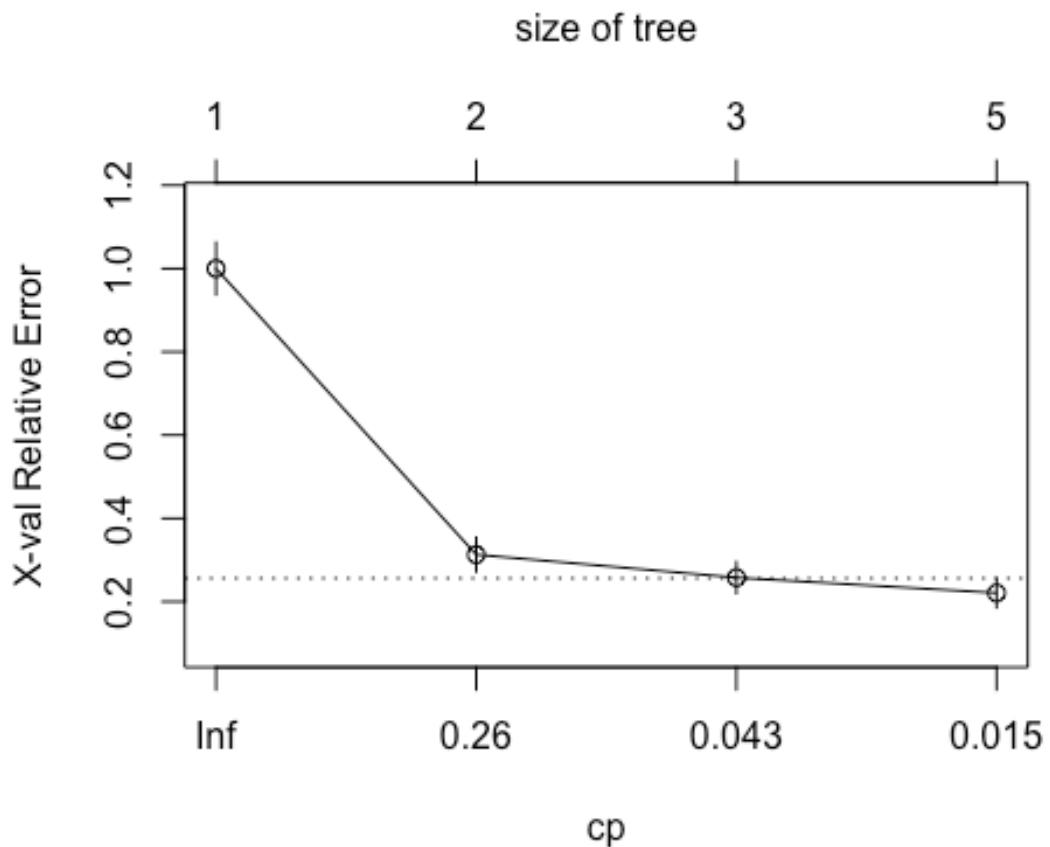
```

```

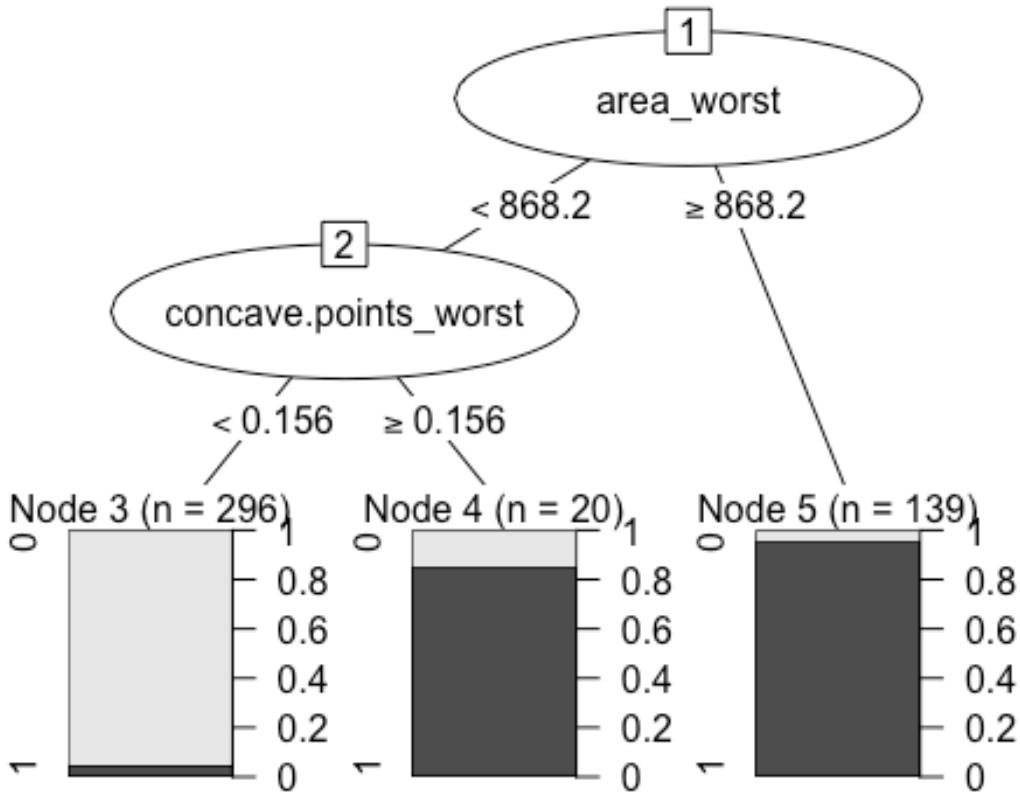
##  compactness_se + concavity_se + concave.points_se + symmetry_se +
##  fractal_dimension_se + radius_worst + texture_worst + perimeter_worst
+
##  area_worst + smoothness_worst + compactness_worst + concavity_worst +
##  concave.points_worst + symmetry_worst + fractal_dimension_worst,
##  data = train_data_tree)
##
## Variables actually used in tree construction:
## [1] area_worst          concave.points_worst texture_mean
##
## Root node error: 163/455 = 0.35824
##
## n= 455
##
##          CP nsplit rel_error xerror      xstd
## 1 0.779141     0 1.000000 1.00000 0.062747
## 2 0.085890     1 0.220859 0.31288 0.041284
## 3 0.021472     2 0.134969 0.25767 0.037880
## 4 0.010000     4 0.092025 0.22086 0.035324

plotcp(tree_1)

```



```
#prune the tree using cp=0.03 from plotcp()
tree_1_prune <- prune(tree_1, cp=0.03)
plot(as.party(tree_1_prune))
```



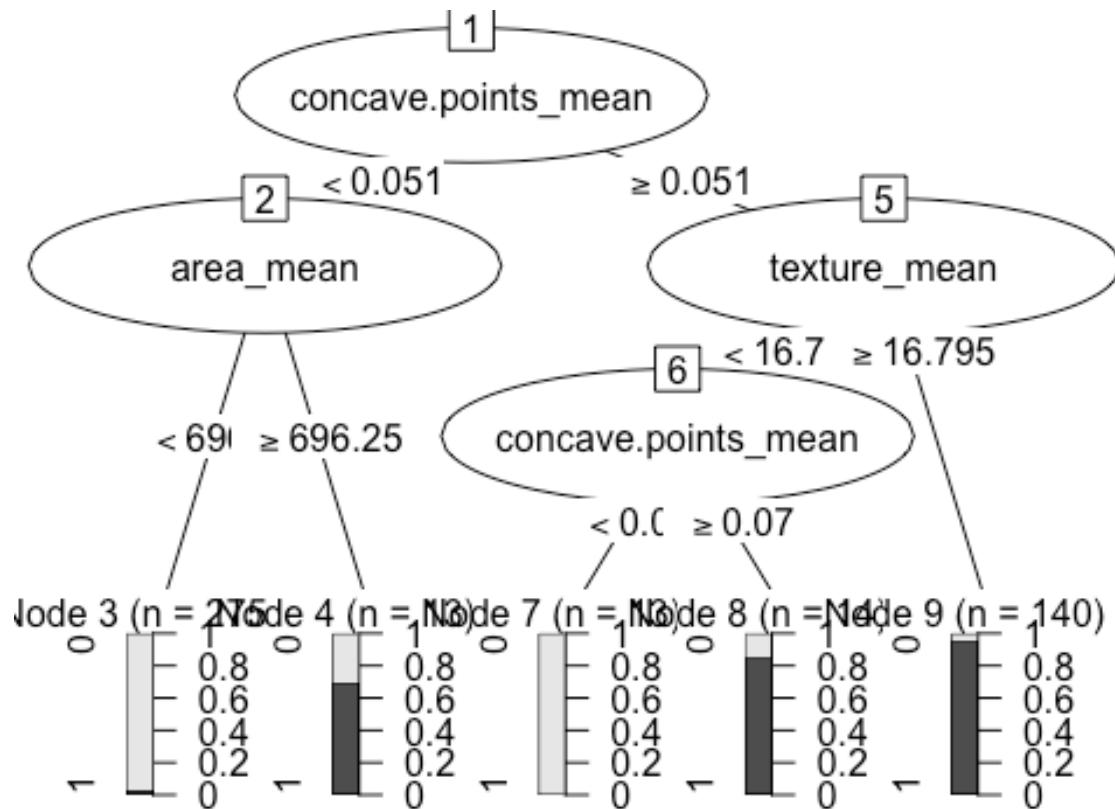
```
tree_1_prune_pred <- predict(tree_1_prune, newdata = test_data_tree, type = "class")
tree_1_prune_pred

## 169 532 219 291 135 371 272 325 153 265 545 200 302 192 177 38 254 210
## 1 0 1 0 1 1 0 0 1 1 0 1 0 1 0 0 1 1
## 168 87 530 521 128 88 533 227 116 432 517 99 428 260 341 509 525 324
## 1 0 0 0 1 1 0 0 0 0 1 0 0 1 0 1 0 1
## 19 490 225 285 95 362 28 7 407 143 262 220 2 339 159 278 338 454
## 1 1 0 0 1 0 1 1 1 0 1 1 1 0 0 1 1 0
## 70 410 498 112 211 231 188 437 16 382 191 226 224 178 280 15 352 385
## 0 0 0 0 1 1 0 0 1 0 1 1 1 1 0 1 1 0
## 94 216 54 290 357 255 438 122 388 504 442 321 62 5 400 238 426 63
## 0 1 1 0 0 1 0 1 0 1 1 0 0 1 0 1 0 1
## 77 349 464 83 217 235 519 37 366 203 60 526 163 118 552 293 283 448
## 0 0 0 1 0 0 0 0 1 1 0 0 1 1 0 0 1 0
## 477 221 33 455 458 347
## 0 0 1 0 0 0
## Levels: 0 1
```

```



```



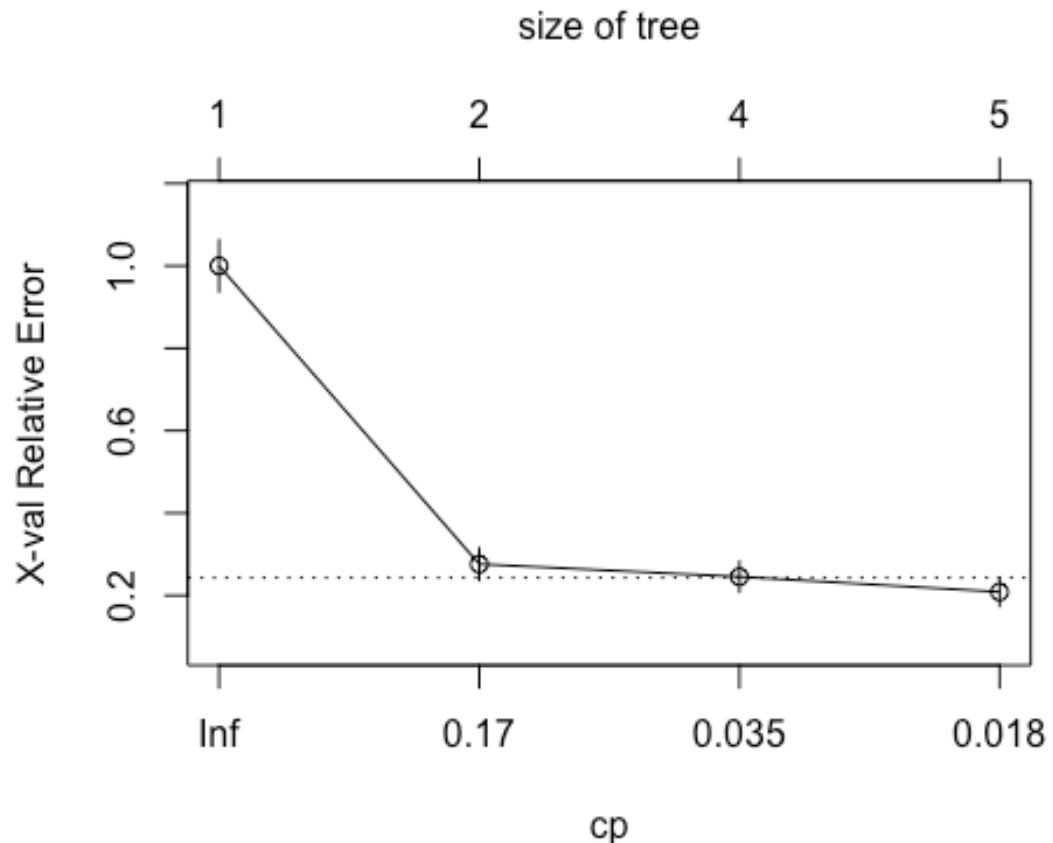
```

#determine complexity parameter for tree
printcp(tree_2)

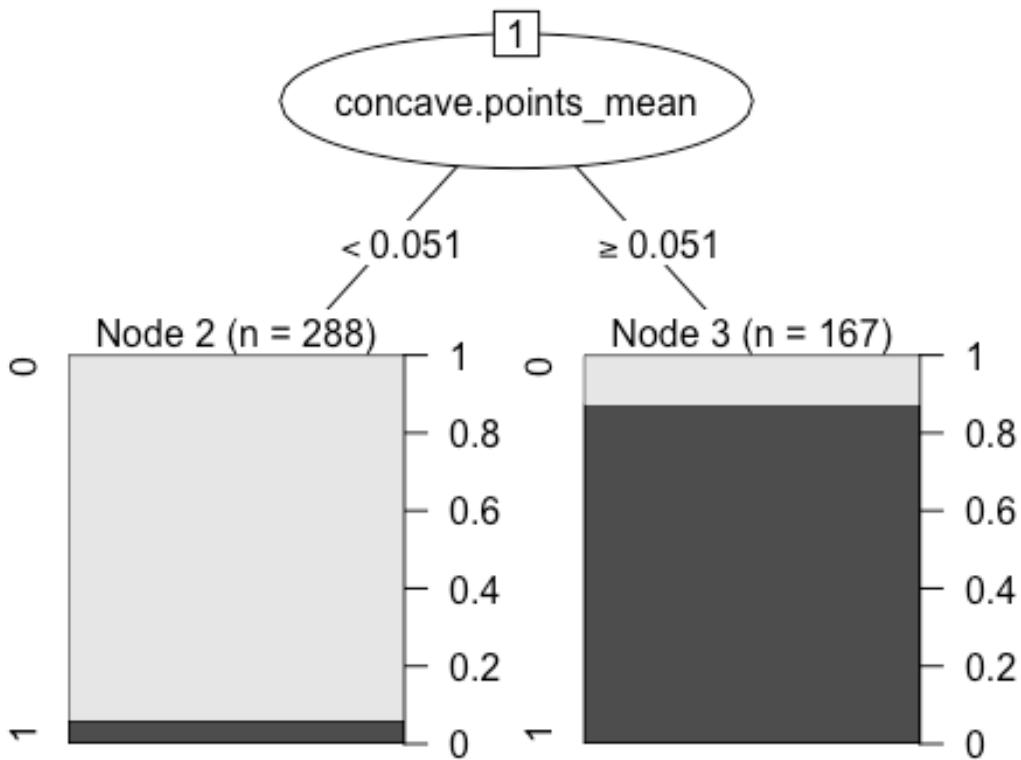
##
## Classification tree:
## rpart(formula = diagnosis ~ radius_mean + texture_mean + perimeter_mean +
##        area_mean + smoothness_mean + compactness_mean + concavity_mean +
##        concave.points_mean + symmetry_mean + fractal_dimension_mean,
##        data = train_data_tree)
##
## Variables actually used in tree construction:
## [1] area_mean          concave.points_mean texture_mean
##
## Root node error: 163/455 = 0.35824
##
## n= 455
##
##          CP nsplit rel error  xerror      xstd
## 1 0.766871      0  1.00000 1.00000  0.062747
## 2 0.039877      1  0.23313 0.27607  0.039067
## 3 0.030675      3  0.15337 0.24540  0.037056
## 4 0.010000      4  0.12270 0.20859  0.034410

```

```
plotcp(tree_2)
```



```
#prune the tree using cp=0.04 from plotcp()
tree_2_prune <- prune(tree_2, cp=0.04)
plot(as.party(tree_2_prune))
```



```

tree_2_prune_pred <- predict(tree_2_prune,newdata = test_data_tree,type = "class")
tree_2_prune_pred

## 169 532 219 291 135 371 272 325 153 265 545 200 302 192 177 38 254 210
## 1 0 1 1 1 1 0 0 1 1 0 1 0 0 0 0 1 0
## 168 87 530 521 128 88 533 227 116 432 517 99 428 260 341 509 525 324
## 1 0 0 0 1 1 0 0 0 0 1 0 0 1 0 0 0 1
## 19 490 225 285 95 362 28 7 407 143 262 220 2 339 159 278 338 454
## 1 0 0 0 1 0 1 1 0 0 0 1 1 0 0 1 1 1
## 70 410 498 112 211 231 188 437 16 382 191 226 224 178 280 15 352 385
## 0 0 0 1 1 1 0 0 1 0 1 0 1 1 0 1 1 0
## 94 216 54 290 357 255 438 122 388 504 442 321 62 5 400 238 426 63
## 0 1 1 0 1 1 0 1 0 1 1 0 0 1 0 1 0 1
## 77 349 464 83 217 235 519 37 366 203 60 526 163 118 552 293 283 448
## 1 0 0 1 0 0 1 1 1 0 0 1 1 0 0 1 0 1
## 477 221 33 455 458 347
## 0 0 1 0 0 0
## Levels: 0 1

table(test_data_tree$diagnosis,tree_2_prune_pred)

##      tree_2_prune_pred
##      0 1

```

```

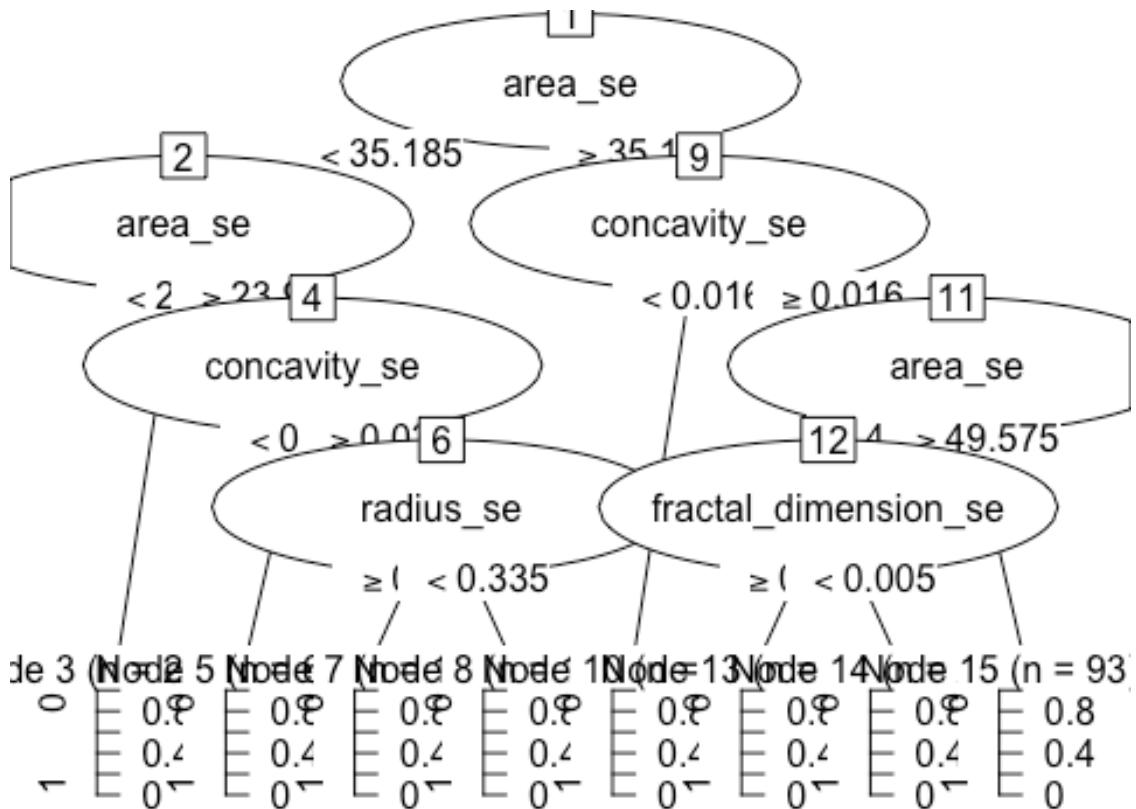
##  0 58  7
##  1 3 46

#decision tree from the training data set predicting
#the diagnosis variable using all standard error variables
tree_3 <- rpart(diagnosis~radius_se+texture_se+perimeter_se+area_se+smoothness_se+
                  compactness_se+concavity_se+concave.points_se+symmetry_se+
                  fractal_dimension_se,
                  data = train_data_tree)
tree_3

## n= 455
##
## node), split, n, loss, yval, (yprob)
##           * denotes terminal node
##
##  1) root 455 163 0 (0.64175824 0.35824176)
##      2) area_se< 35.185 315 39 0 (0.87619048 0.12380952)
##          4) area_se< 23.93 224 14 0 (0.93750000 0.06250000) *
##          5) area_se>=23.93 91 25 0 (0.72527473 0.27472527)
##          10) concavity_se< 0.033165 60 6 0 (0.90000000 0.10000000) *
##          11) concavity_se>=0.033165 31 12 1 (0.38709677 0.61290323)
##              22) radius_se>=0.33495 18 7 0 (0.61111111 0.38888889) *
##              23) radius_se< 0.33495 13 1 1 (0.07692308 0.92307692) *
##          3) area_se>=35.185 140 16 1 (0.11428571 0.88571429)
##              6) concavity_se< 0.01568 8 2 0 (0.75000000 0.25000000) *
##              7) concavity_se>=0.01568 132 10 1 (0.07575758 0.92424242)
##                  14) area_se< 49.575 39 10 1 (0.25641026 0.74358974)
##                      28) fractal_dimension_se>=0.0046175 10 3 0 (0.70000000 0.3000000
0) *
##              29) fractal_dimension_se< 0.0046175 29 3 1 (0.10344828 0.8965517
2) *
##              15) area_se>=49.575 93 0 1 (0.00000000 1.00000000) *

#partykit based plot of the tree produced
plot(as.party(tree_3))

```

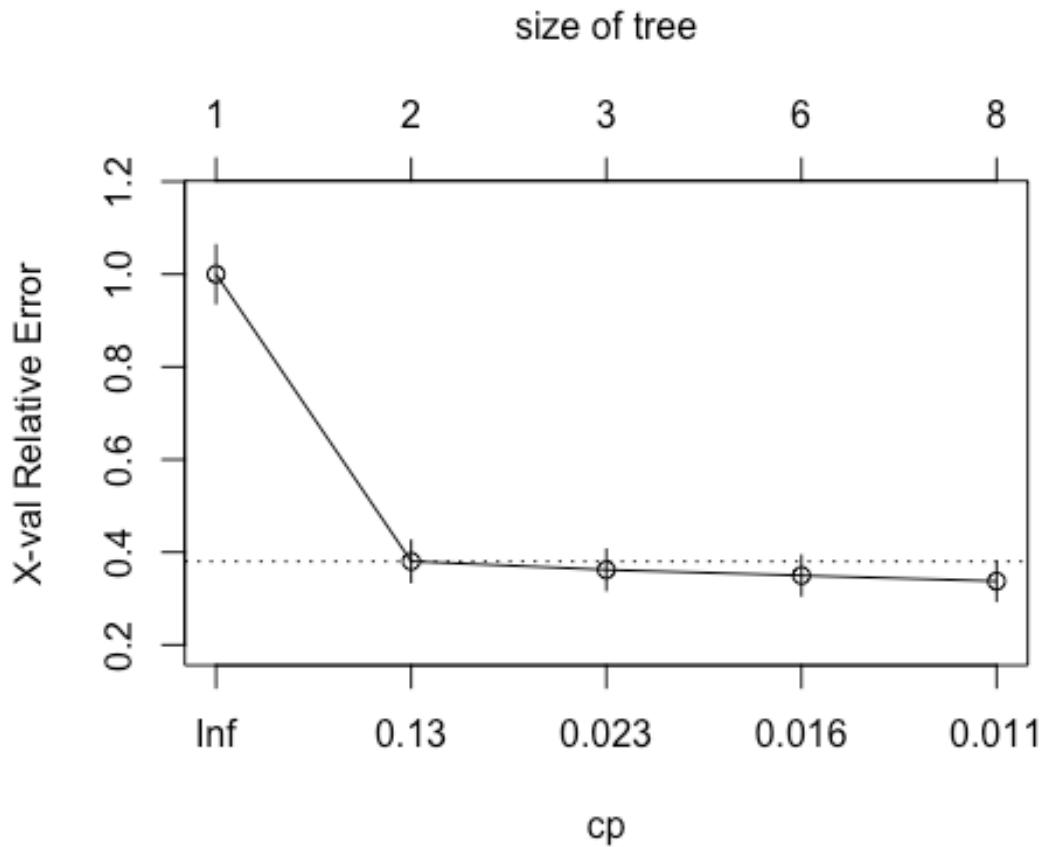


```
#determine complexity parameter for tree
printcp(tree_3)

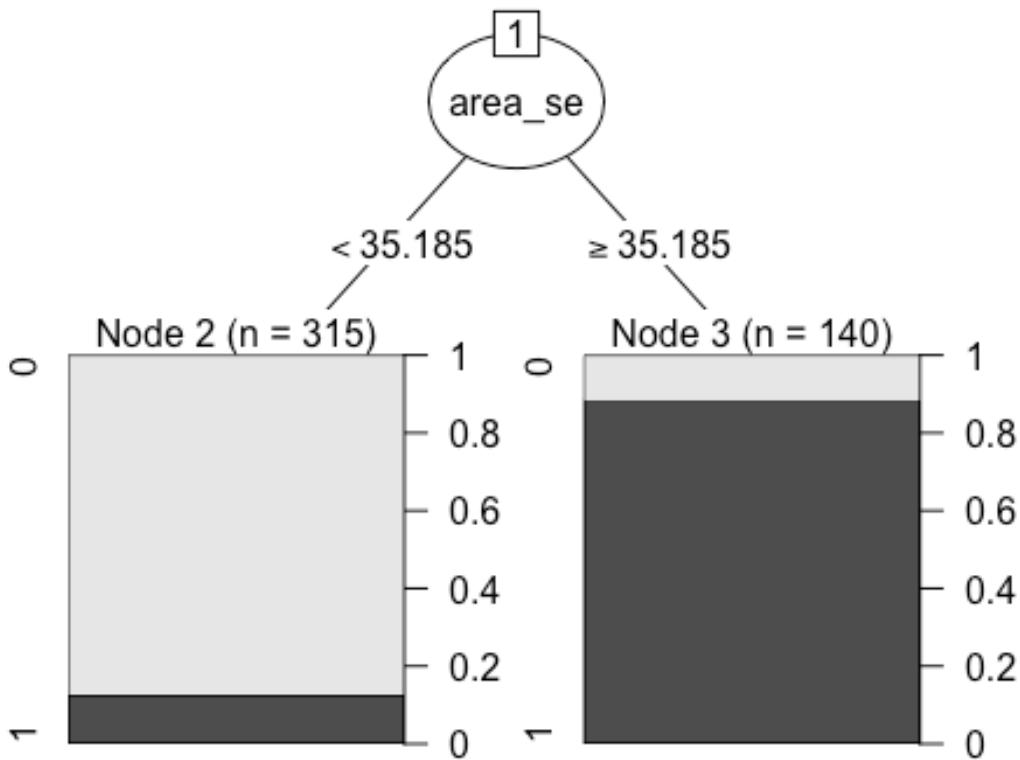
##
## Classification tree:
## rpart(formula = diagnosis ~ radius_se + texture_se + perimeter_se +
##        area_se + smoothness_se + compactness_se + concavity_se +
##        concave.points_se + symmetry_se + fractal_dimension_se, data = train_d
ata_tree)
##
## Variables actually used in tree construction:
## [1] area_se           concavity_se           fractal_dimension_se
## [4] radius_se
##
## Root node error: 163/455 = 0.35824
##
## n= 455
##
##          CP nsplit rel error  xerror      xstd
## 1 0.662577      0  1.00000 1.00000 0.062747
## 2 0.024540      1  0.33742 0.38037 0.044895
## 3 0.021472      2  0.31288 0.36196 0.043962
```

```
## 4 0.012270      5  0.24540 0.34969 0.043320
## 5 0.010000      7  0.22086 0.33742 0.042660

plotcp(tree_3)
```



```
#prune the tree using cp=0.04 from plotcp()
tree_3_prune <- prune(tree_3, cp=0.04)
plot(as.party(tree_3_prune))
```



```

tree_3_prune_pred <- predict(tree_3_prune,newdata = test_data_tree,type = "class")
tree_3_prune_pred

## 169 532 219 291 135 371 272 325 153 265 545 200 302 192 177 38 254 210
## 1 0 1 1 1 1 0 0 1 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## 168 87 530 521 128 88 533 227 116 432 517 99 428 260 341 509 525 324
## 1 1 0 0 1 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
## 19 490 225 285 95 362 28 7 407 143 262 220 2 339 159 278 338 454
## 1 0 0 0 1 0 1 1 0 0 1 1 1 1 1 0 0 0 1 1 0 0 0 0 1 1 0 0 0 0 1 1 0 0 0 0 1 0
## 70 410 498 112 211 231 188 437 16 382 191 226 224 178 280 15 352 385
## 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0
## 94 216 54 290 357 255 438 122 388 504 442 321 62 5 400 238 426 63
## 0 0 1 0 0 1 0 1 0 1 1 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 1 0 1
## 77 349 464 83 217 235 519 37 366 203 60 526 163 118 552 293 283 448
## 0 0 0 1 0 0 0 0 1 1 0 0 0 1 1 0 0 1 1 0 0 1 1 0 0 0 0 1 0 0 1 0 0 1 0
## 477 221 33 455 458 347
## 0 0 1 0 0 0
## Levels: 0 1

```

```
table(test_data_tree$diagnosis,tree_3_prune_pred)
```

```
##      tree_3_prune_pred
##      0 1
```

```

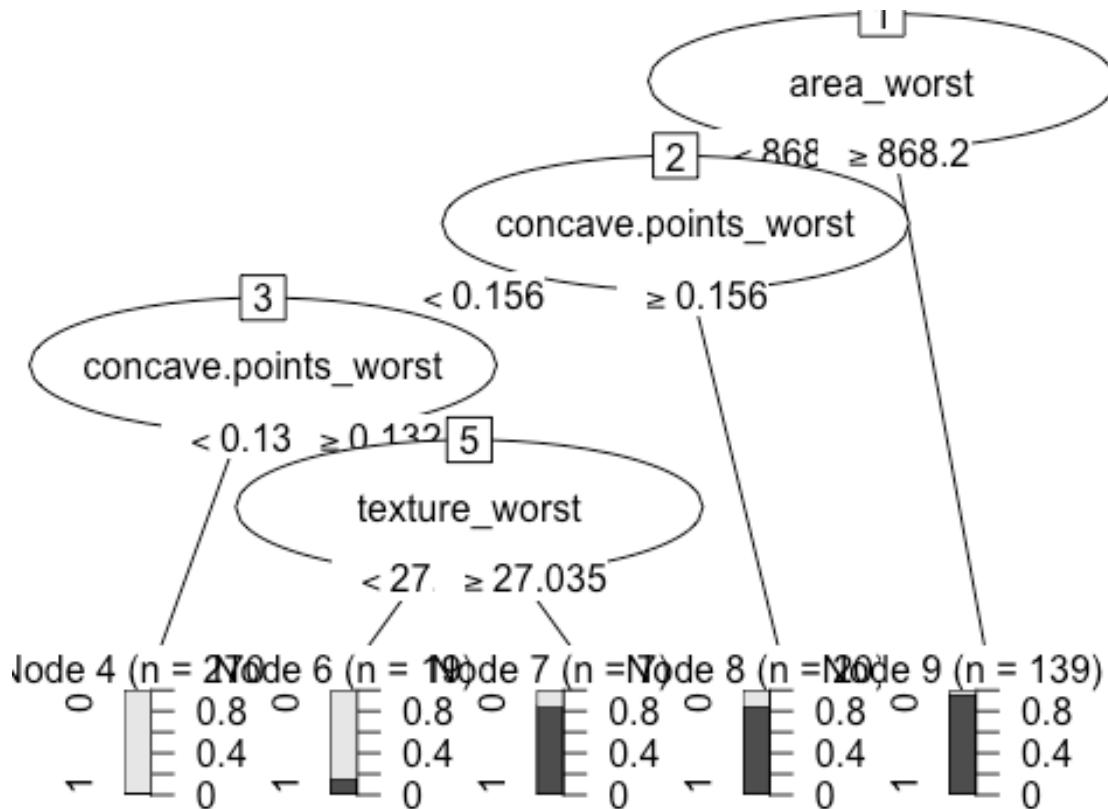
##  0 61  4
##  1 12 37

#decision tree from the training data set predicting
#the diagnosis variable using all standard error variables
tree_4 <- rpart(diagnosis~ radius_worst+texture_worst+perimeter_worst+area_wor
rst+smoothness_worst+
    compactness_worst+concavity_worst+concave.points_worst+symm
etry_worst+
    fractal_dimension_worst,
    data=train_data_tree)
tree_4

## n= 455
##
## node), split, n, loss, yval, (yprob)
##      * denotes terminal node
##
##  1) root 455 163 0 (0.64175824 0.35824176)
##      2) area_worst< 868.2 316 30 0 (0.90506329 0.09493671)
##          4) concave.points_worst< 0.1563 296 13 0 (0.95608108 0.04391892)
##          8) concave.points_worst< 0.13235 270 4 0 (0.98518519 0.01481481)
##
##          9) concave.points_worst>=0.13235 26 9 0 (0.65384615 0.34615385)
##              18) texture_worst< 27.035 19 3 0 (0.84210526 0.15789474) *
##                  19) texture_worst>=27.035 7 1 1 (0.14285714 0.85714286) *
##              5) concave.points_worst>=0.1563 20 3 1 (0.15000000 0.85000000) *
##              3) area_worst>=868.2 139 6 1 (0.04316547 0.95683453) *

#partykit based plot of the tree produced
plot(as.party(tree_4))

```

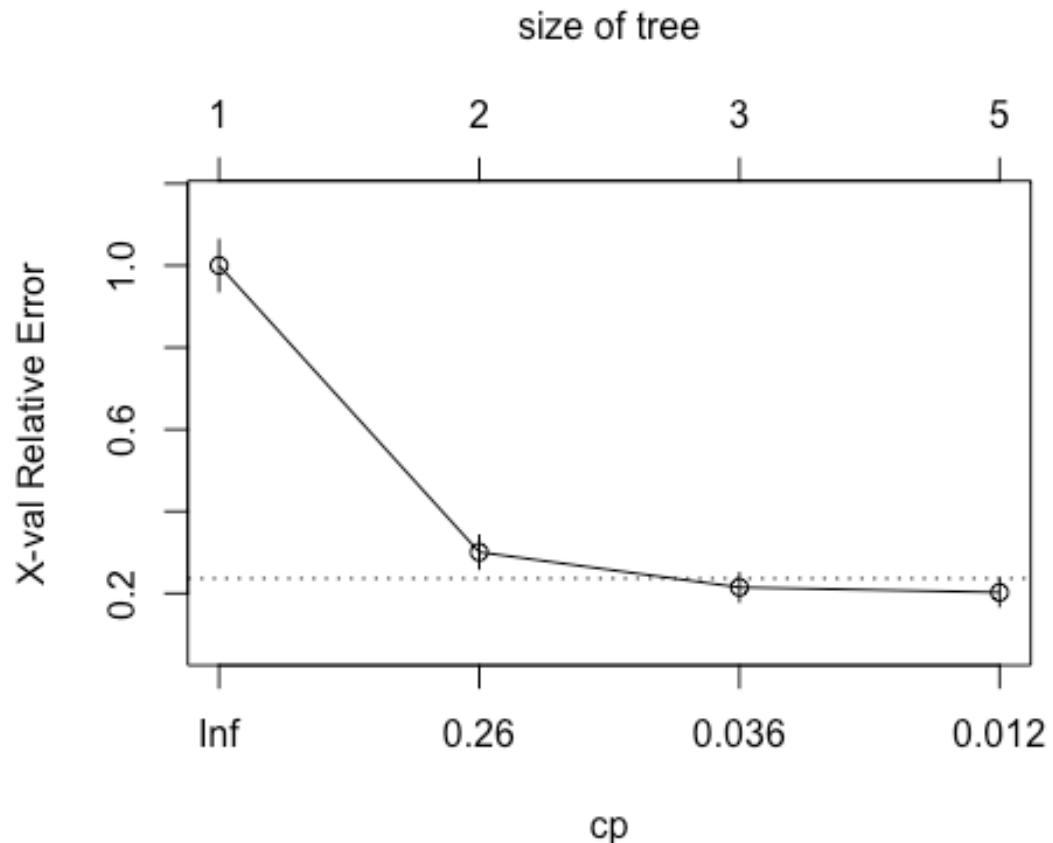


```

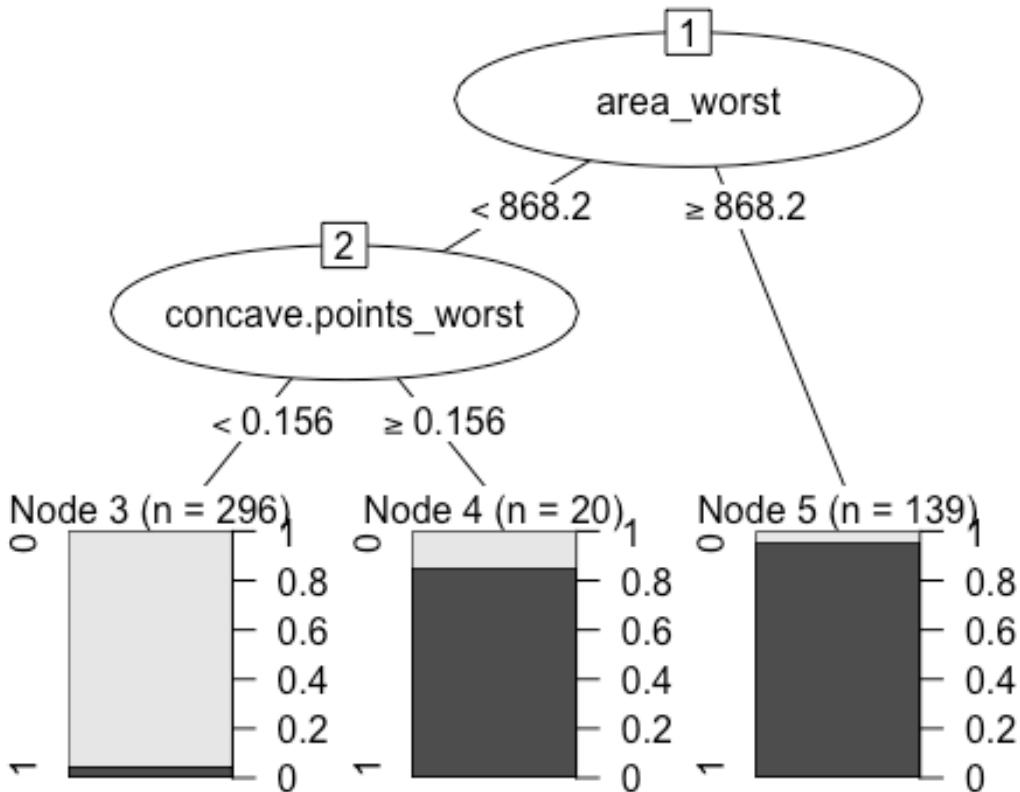
#determine complexity parameter for tree
printcp(tree_4)

##
## Classification tree:
## rpart(formula = diagnosis ~ radius_worst + texture_worst + perimeter_worst +
##        area_worst + smoothness_worst + compactness_worst + concavity_worst +
##        concave.points_worst + symmetry_worst + fractal_dimension_worst,
##        data = train_data_tree)
##
## Variables actually used in tree construction:
## [1] area_worst          concave.points_worst texture_worst
##
## Root node error: 163/455 = 0.35824
##
## n= 455
##
##          CP nsplit rel error xerror      xstd
## 1 0.779141      0  1.00000 1.00000 0.062747
## 2 0.085890      1  0.22086 0.30061 0.040567
## 3 0.015337      2  0.13497 0.21472 0.034871
## 4 0.010000      4  0.10429 0.20245 0.033941
  
```

```
plotcp(tree_4)
```



```
#prune the tree using cp=0.04 from plotcp()
tree_4_prune <- prune(tree_4, cp=0.04)
plot(as.party(tree_4_prune))
```



```

tree_4_prune_pred <- predict(tree_4_prune,newdata = test_data_tree,type = "class")
tree_4_prune_pred

## 169 532 219 291 135 371 272 325 153 265 545 200 302 192 177 38 254 210
## 1 0 1 0 1 1 0 0 1 1 0 1 0 1 0 0 1 1
## 168 87 530 521 128 88 533 227 116 432 517 99 428 260 341 509 525 324
## 1 0 0 0 1 1 0 0 0 0 1 0 0 1 0 1 0 1
## 19 490 225 285 95 362 28 7 407 143 262 220 2 339 159 278 338 454
## 1 1 0 0 1 0 1 1 1 0 1 1 1 0 0 1 1 0
## 70 410 498 112 211 231 188 437 16 382 191 226 224 178 280 15 352 385
## 0 0 0 0 1 1 0 0 1 0 1 1 1 1 0 0 1 1 0
## 94 216 54 290 357 255 438 122 388 504 442 321 62 5 400 238 426 63
## 0 1 1 0 0 1 0 1 0 1 1 0 0 1 0 1 0 1
## 77 349 464 83 217 235 519 37 366 203 60 526 163 118 552 293 283 448
## 0 0 0 1 0 0 0 0 1 1 0 0 1 1 0 0 1 0
## 477 221 33 455 458 347
## 0 0 1 0 0 0
## Levels: 0 1

table(test_data_tree$diagnosis,tree_4_prune_pred)

##      tree_4_prune_pred
##      0 1
  
```

```

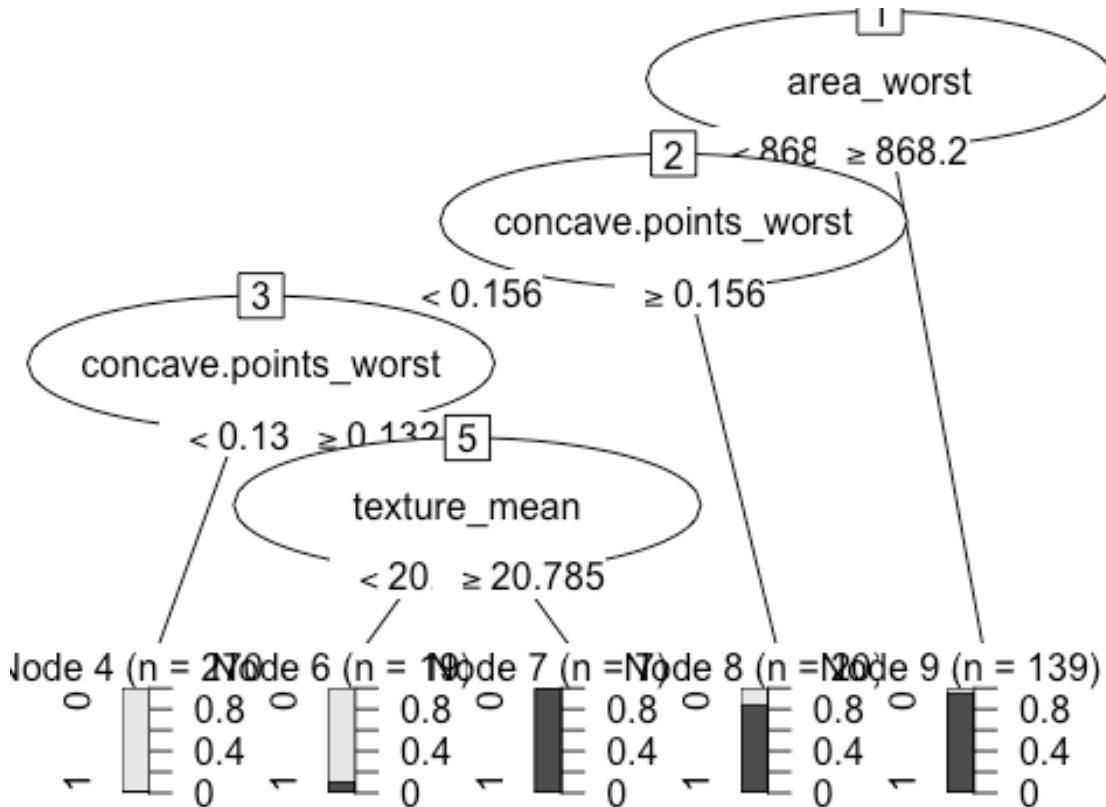
##  0 60  5
##  1  2 47

#decision tree from the training data set predicting
#the diagnosis variable using the highlighted variables from the first 4 tree
s
tree_5 <- rpart(diagnosis~ concave.points_mean+area_mean+texture_mean+perimeter_mean+
                  area_se+symmetry_se+concavity_se+radius_se+fractal_dimension+
n_se+
                  area_worst+concave.points_worst+concavity_worst+texture_wor
st,
                  data=train_data_tree)
tree_5

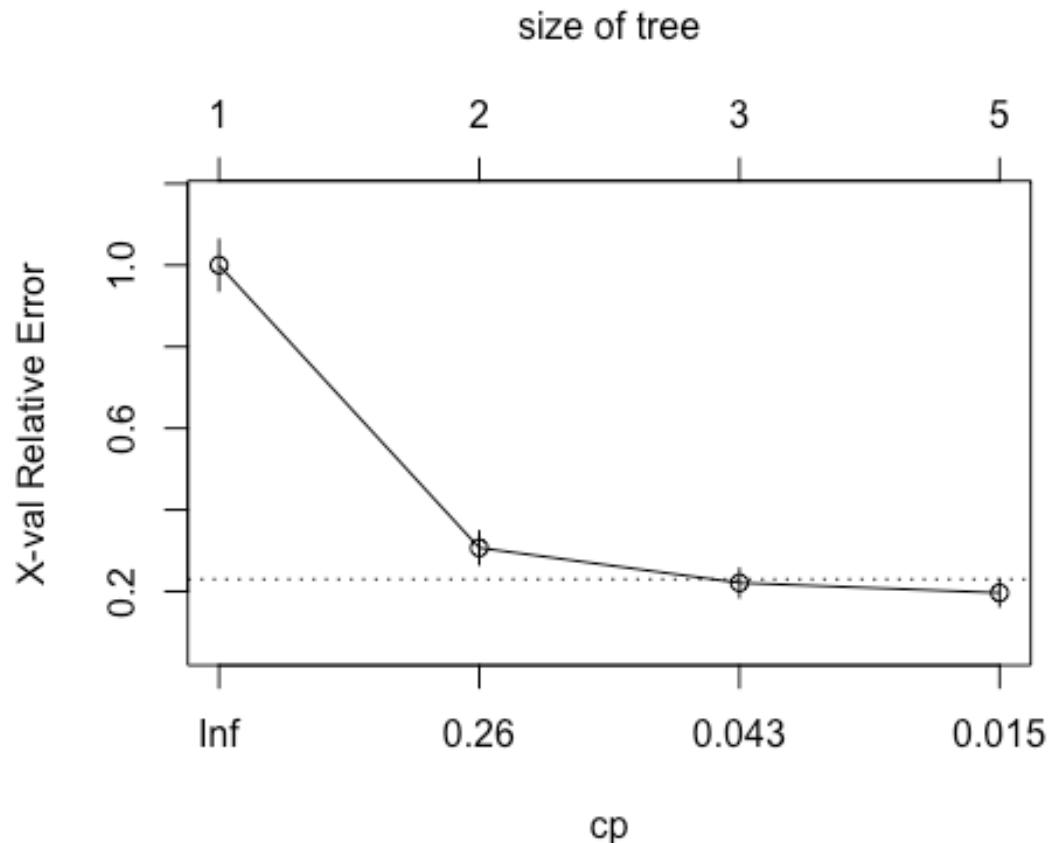
## n= 455
##
## node), split, n, loss, yval, (yprob)
##      * denotes terminal node
##
##  1) root 455 163 0 (0.64175824 0.35824176)
##      2) area_worst< 868.2 316 30 0 (0.90506329 0.09493671)
##          4) concave.points_worst< 0.1563 296 13 0 (0.95608108 0.04391892)
##              8) concave.points_worst< 0.13235 270 4 0 (0.98518519 0.01481481)
##
##              9) concave.points_worst>=0.13235 26 9 0 (0.65384615 0.34615385)
##                  18) texture_mean< 20.785 19 2 0 (0.89473684 0.10526316) *
##                      19) texture_mean>=20.785 7 0 1 (0.00000000 1.00000000) *
##              5) concave.points_worst>=0.1563 20 3 1 (0.15000000 0.85000000) *
##              3) area_worst>=868.2 139 6 1 (0.04316547 0.95683453) *

#partykit based plot of the tree produced
plot(as.party(tree_5))

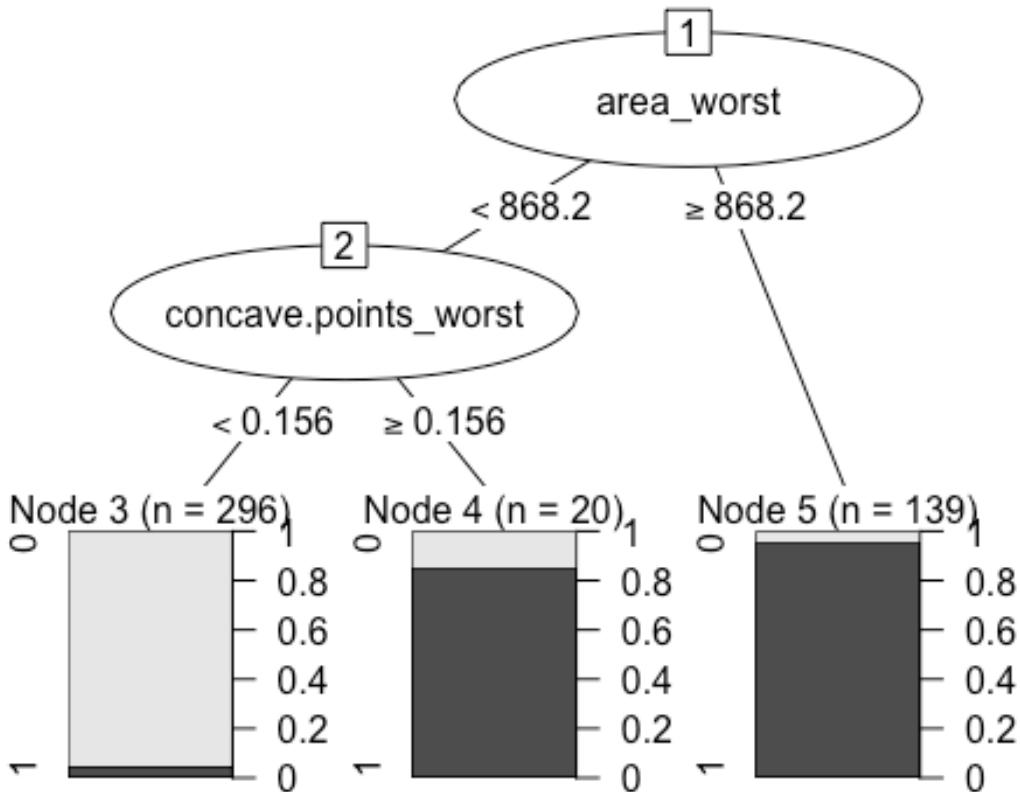
```



```
plotcp(tree_5)
```



```
#prune the tree using cp=0.05 from plotcp()
tree_5_prune <- prune(tree_5, cp=0.05)
plot(as.party(tree_5_prune))
```



```

tree_5_prune_pred <- predict(tree_5_prune, newdata = test_data_tree, type = "class")
tree_5_prune_pred

## 169 532 219 291 135 371 272 325 153 265 545 200 302 192 177 38 254 210
## 1 0 1 0 1 1 0 0 1 1 0 1 0 0 0 0 1 1 1
## 168 87 530 521 128 88 533 227 116 432 517 99 428 260 341 509 525 324
## 1 0 0 0 1 1 0 0 0 0 1 0 0 1 0 1 0 1
## 19 490 225 285 95 362 28 7 407 143 262 220 2 339 159 278 338 454
## 1 1 0 0 1 0 1 1 1 0 1 1 1 0 0 1 1 0
## 70 410 498 112 211 231 188 437 16 382 191 226 224 178 280 15 352 385
## 0 0 0 0 1 1 0 0 1 0 1 1 1 1 0 0 1 1 0
## 94 216 54 290 357 255 438 122 388 504 442 321 62 5 400 238 426 63
## 0 1 1 0 0 1 0 1 0 1 1 0 0 1 0 1 0 1
## 77 349 464 83 217 235 519 37 366 203 60 526 163 118 552 293 283 448
## 0 0 0 1 0 0 0 0 1 1 0 0 1 1 0 0 1 0
## 477 221 33 455 458 347
## 0 0 1 0 0 0
## Levels: 0 1

table(test_data_tree$diagnosis, tree_5_prune_pred)

##      tree_5_prune_pred
##      0 1
  
```

```

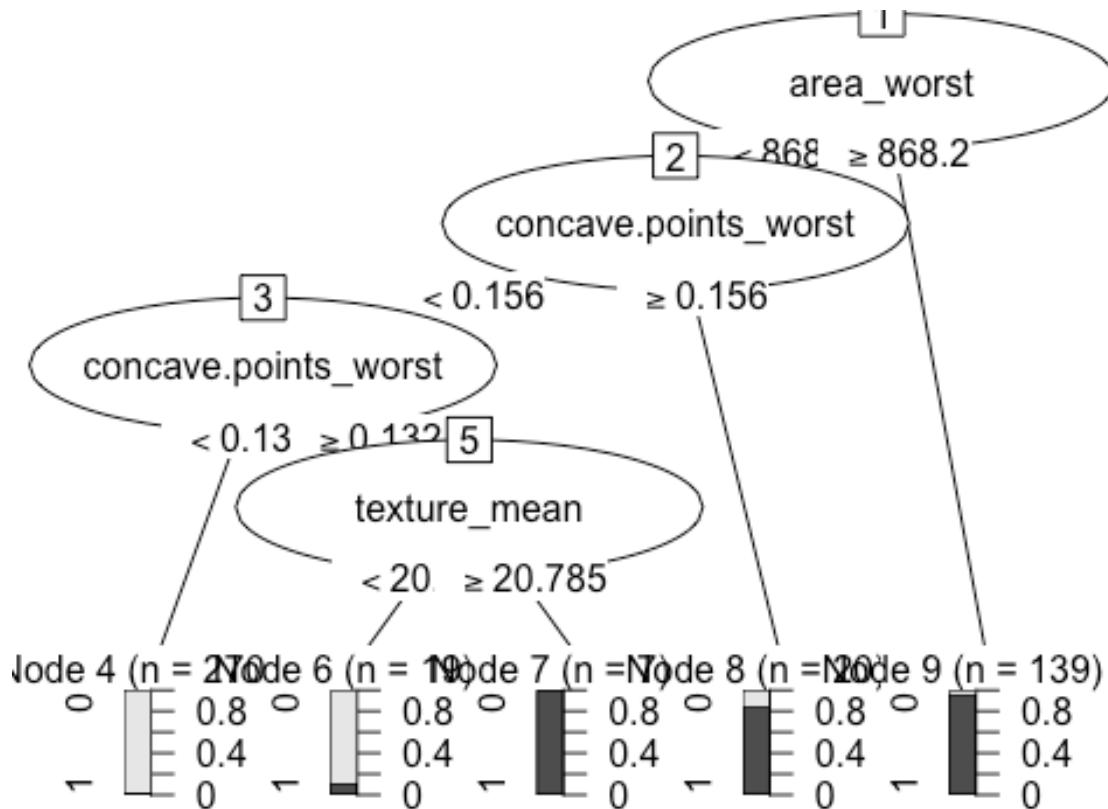
##   0 60  5
##   1  2 47

#decision tree from the training data set predicting
#the diagnosis variable using the variables area_worst,
#concave.points_worst, concavity_worst, and texture_mean
#highlighted in tree_5
tree_6 <- rpart(diagnosis~area_worst+concave.points_worst+concavity_worst+texture_mean,
                  data=train_data_tree)
tree_6

## n= 455
##
## node), split, n, loss, yval, (yprob)
##           * denotes terminal node
##
## 1) root 455 163 0 (0.64175824 0.35824176)
##    2) area_worst< 868.2 316 30 0 (0.90506329 0.09493671)
##      4) concave.points_worst< 0.1563 296 13 0 (0.95608108 0.04391892)
##        8) concave.points_worst< 0.13235 270 4 0 (0.98518519 0.01481481)
##
##          9) concave.points_worst>=0.13235 26 9 0 (0.65384615 0.34615385)
##            18) texture_mean< 20.785 19 2 0 (0.89473684 0.10526316) *
##              19) texture_mean>=20.785 7 0 1 (0.00000000 1.00000000) *
##            5) concave.points_worst>=0.1563 20 3 1 (0.15000000 0.85000000) *
##        3) area_worst>=868.2 139 6 1 (0.04316547 0.95683453) *

#partykit based plot of the tree produced
plot(as.party(tree_6))

```

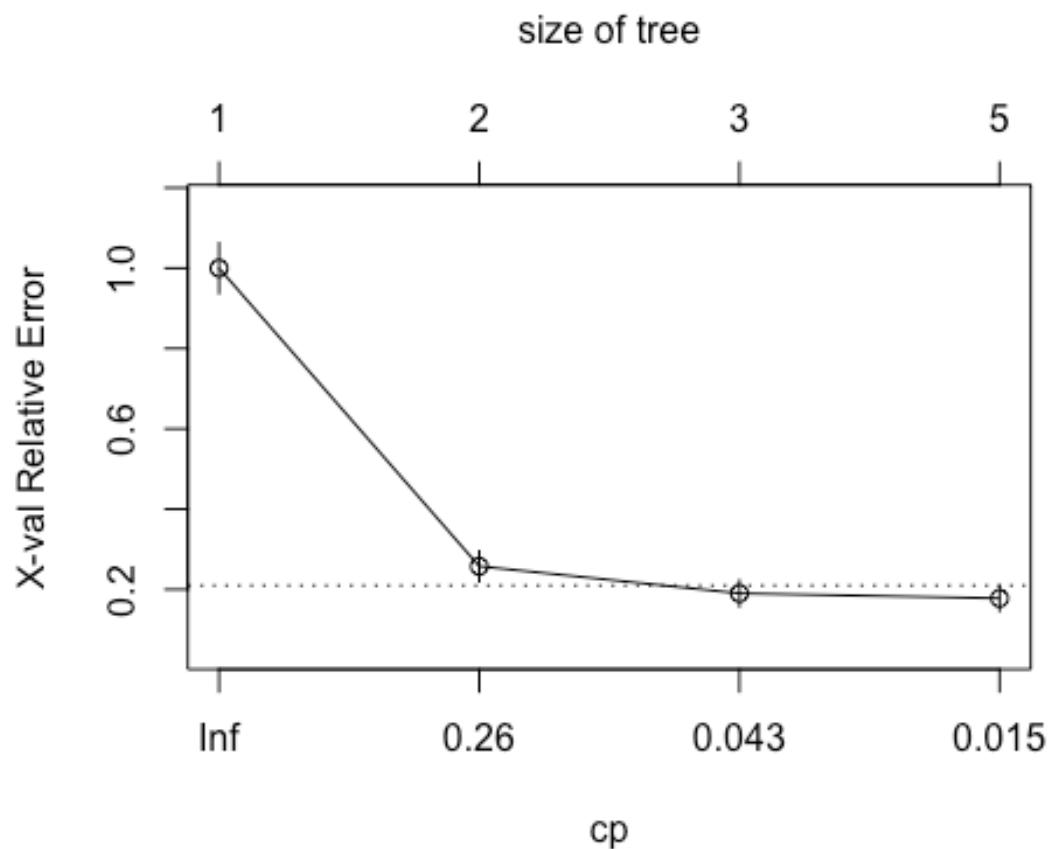


```

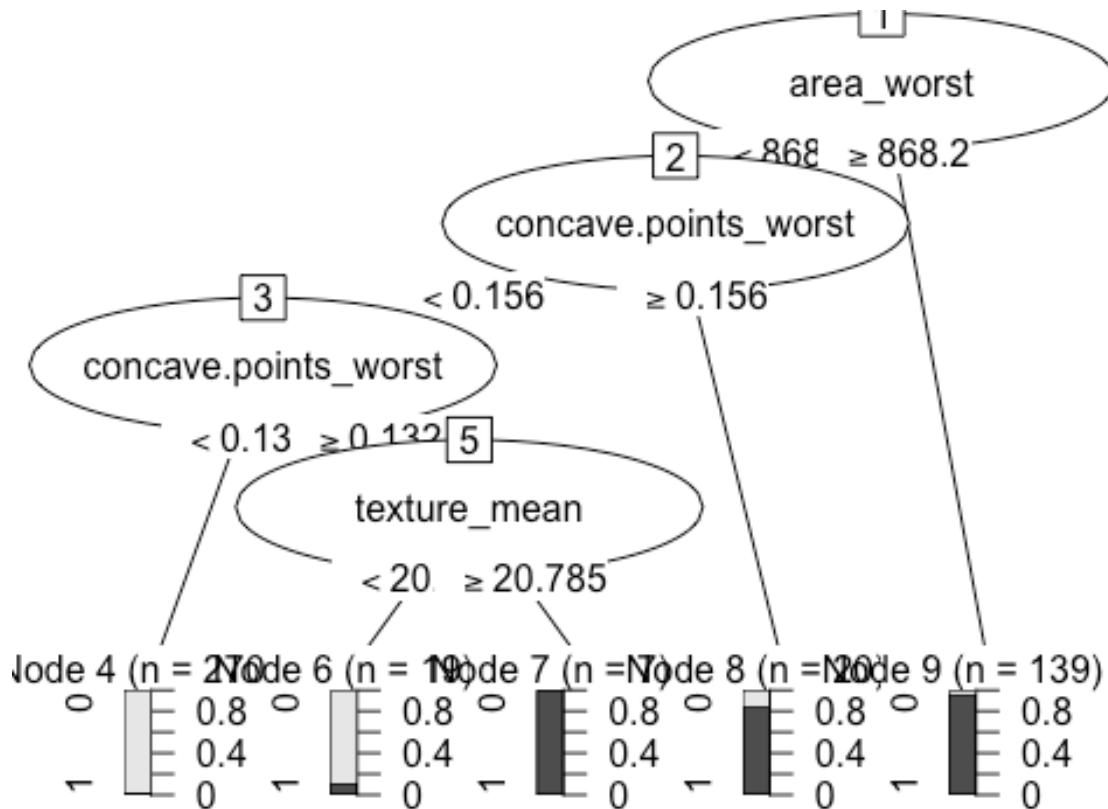
#determine complexity parameter for tree
printcp(tree_6)

##
## Classification tree:
## rpart(formula = diagnosis ~ area_worst + concave.points_worst +
##       concavity_worst + texture_mean, data = train_data_tree)
##
## Variables actually used in tree construction:
## [1] area_worst      concave.points_worst texture_mean
##
## Root node error: 163/455 = 0.35824
##
## n= 455
##
##          CP nsplit rel error  xerror      xstd
## 1 0.779141      0  1.00000 1.00000 0.062747
## 2 0.085890      1  0.220859 0.25767 0.037880
## 3 0.021472      2  0.134969 0.19018 0.032974
## 4 0.010000      4  0.092025 0.17791 0.031968

plotcp(tree_6)
  
```



```
#prune the tree using cp=0.02 from plotcp()
tree_6_prune <- prune(tree_6, cp=0.02)
plot(as.party(tree_6_prune))
```



```

tree_6_prune_pred <- predict(tree_6_prune,newdata = test_data_tree,type = "class")
tree_6_prune_pred

## 169 532 219 291 135 371 272 325 153 265 545 200 302 192 177 38 254 210
## 1 0 1 0 1 1 0 0 1 1 0 1 0 1 0 0 0 1 1 0 1 0 0 1 1 0 1 1 0 1 1 0 1 1 0 1 1
## 168 87 530 521 128 88 533 227 116 432 517 99 428 260 341 509 525 324
## 1 0 0 0 1 1 0 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1
## 19 490 225 285 95 362 28 7 407 143 262 220 2 339 159 278 338 454
## 1 1 0 0 1 0 1 1 1 0 1 1 1 0 1 1 1 0 0 1 1 0 0 1 1 1 0 0 1 1 1 0 1 1 0
## 70 410 498 112 211 231 188 437 16 382 191 226 224 178 280 15 352 385
## 0 0 0 0 1 1 0 0 1 0 1 1 1 0 1 1 1 1 0 1 1 0 0 1 1 1 0 1 1 1 0 1 1 0
## 94 216 54 290 357 255 438 122 388 504 442 321 62 5 400 238 426 63
## 0 1 1 0 0 1 0 1 0 1 1 1 0 0 1 0 1 0 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1
## 77 349 464 83 217 235 519 37 366 203 60 526 163 118 552 293 283 448
## 0 0 0 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 1 1 1 0 0 1 0 0 1 0 1 0 0 1 0 1 0
## 477 221 33 455 458 347
## 0 0 1 0 0 0
## Levels: 0 1

table(test_data_tree$diagnosis,tree_6_prune_pred)

##      tree_6_prune_pred
##      0 1
  
```

```

## 0 60 5
## 1 1 48

#5 - Build a classification algorithm using random forests/bagging.
#Adjust the parameters of the forest appropriately.

#Create a copy of the dataset for use in the random forests/bagging analysis
cancer_bag_all <- cancer
#make the diagnosis variable binary, malignant = 1, benign = 0
#cancer_bag_all$diagnosis <- ifelse(cancer_bag_all$diagnosis=='M',1,0)

#create the rescale function
rescale_x <- function(x){(x-min(x))/(max(x)-min(x))}

#use the rescale function to rescale the numeric observations
cancer_bag_all$radius_mean <- rescale_x(cancer_bag_all$radius_mean)
cancer_bag_all$texture_mean <- rescale_x(cancer_bag_all$texture_mean)
cancer_bag_all$perimeter_mean <- rescale_x(cancer_bag_all$perimeter_mean)
cancer_bag_all$area_mean <- rescale_x(cancer_bag_all$area_mean)
cancer_bag_all$smoothness_mean <- rescale_x(cancer_bag_all$smoothness_mean)
cancer_bag_all$compactness_mean <- rescale_x(cancer_bag_all$compactness_mean)
cancer_bag_all$concavity_mean <- rescale_x(cancer_bag_all$concavity_mean)
cancer_bag_all$concave.points_mean <- rescale_x(cancer_bag_all$concave.points_mean)
cancer_bag_all$symmetry_mean <- rescale_x(cancer_bag_all$symmetry_mean)
cancer_bag_all$fractal_dimension_mean <- rescale_x(cancer_bag_all$fractal_dimension_mean)

cancer_bag_all$radius_worst <- rescale_x(cancer_bag_all$radius_worst)
cancer_bag_all$texture_worst <- rescale_x(cancer_bag_all$texture_worst)
cancer_bag_all$perimeter_worst <- rescale_x(cancer_bag_all$perimeter_worst)
cancer_bag_all$area_worst <- rescale_x(cancer_bag_all$area_worst)
cancer_bag_all$smoothness_worst <- rescale_x(cancer_bag_all$smoothness_worst)
cancer_bag_all$compactness_worst <- rescale_x(cancer_bag_all$compactness_worst)
cancer_bag_all$concavity_worst <- rescale_x(cancer_bag_all$concavity_worst)
cancer_bag_all$concave.points_worst <- rescale_x(cancer_bag_all$concave.points_worst)
cancer_bag_all$symmetry_worst <- rescale_x(cancer_bag_all$symmetry_worst)
cancer_bag_all$fractal_dimension_worst <- rescale_x(cancer_bag_all$fractal_dimension_worst)

cancer_bag_all$radius_se <- rescale_x(cancer_bag_all$radius_se)
cancer_bag_all$texture_se <- rescale_x(cancer_bag_all$texture_se)
cancer_bag_all$perimeter_se <- rescale_x(cancer_bag_all$perimeter_se)
cancer_bag_all$area_se <- rescale_x(cancer_bag_all$area_se)
cancer_bag_all$smoothness_se <- rescale_x(cancer_bag_all$smoothness_se)
cancer_bag_all$compactness_se <- rescale_x(cancer_bag_all$compactness_se)
cancer_bag_all$concavity_se <- rescale_x(cancer_bag_all$concavity_se)
cancer_bag_all$concave.points_se <- rescale_x(cancer_bag_all$concave.points_se)

```

```

e)
cancer_bag_all$symmetry_se <- rescale_x(cancer_bag_all$symmetry_se)
cancer_bag_all$fractal_dimension_se <- rescale_x(cancer_bag_all$fractal_dimension_se)

#Random forest analysis using all variables in the original data set

#set the seed
set.seed(1842)

#save the number of rows in data as n
n_bag_all <- nrow(cancer_bag_all)

#create the test_data dataset of 20% of the data number of rows
test_index_bag_all <- sample.int(n_bag_all, size = round(0.2*n_bag_all))
test_data_bag_all <- cancer_bag_all[test_index_bag_all,]

#create the train_data dataset as the remaining 80%
#number of rows
train_data_bag_all <- cancer_bag_all[-test_index_bag_all,]

#glimpse of both datasets
glimpse(test_data_bag_all)

## Observations: 114
## Variables: 33
## $ id                               <int> 8712766, 91903901, 8811842, 89143602, ...
## $ diagnosis                         <fct> M, B, M, B, M, M, B, B, B, M, B, M, B, ...
## $ radius_mean                       <dbl> 0.4964267, 0.2219225, 0.6067017, 0.351...
## $ texture_mean                      <dbl> 0.5062563, 0.3486642, 0.4007440, 0.338...
## $ perimeter_mean                   <dbl> 0.4996890, 0.2171239, 0.5936701, 0.360...
## $ area_mean                          <dbl> 0.35677625, 0.11567338, 0.46086957, 0.0...
## $ smoothness_mean                  <dbl> 0.4718787, 0.4420872, 0.3719419, 0.315...
## $ compactness_mean                 <dbl> 0.43224342, 0.23050733, 0.34114472, 0.0...
## $ concavity_mean                  <dbl> 0.50585754, 0.09840675, 0.29803187, 0.0...
## $ concave.points_mean             <dbl> 0.51838966, 0.10720676, 0.43195825, 0.0...
## $ symmetry_mean                   <dbl> 0.2414141, 0.4035354, 0.5222222, 0.330...
## $ fractal_dimension_mean          <dbl> 0.28833193, 0.30855097, 0.12320977, 0.0...
## $ radius_se                         <dbl> 0.35361217, 0.03447402, 0.30555857, 0.0...
## $ texture_se                        <dbl> 0.232010255, 0.113662482, 0.182505304, ...
## $ perimeter_se                     <dbl> 0.31004099, 0.02996749, 0.26999011, 0.0...
## $ area_se                           <dbl> 0.21572363, 0.01594702, 0.21964595, 0.0...
## $ smoothness_se                    <dbl> 0.15164701, 0.12027059, 0.17306319, 0.0...
## $ compactness_se                  <dbl> 0.25601586, 0.11279178, 0.22109232, 0.0...
## $ concavity_se                    <dbl> 0.11727273, 0.04646465, 0.08702020, 0.0...
## $ concave.points_se               <dbl> 0.29721538, 0.10035992, 0.32430385, 0.0...
## $ symmetry_se                      <dbl> 0.05020544, 0.09298137, 0.15601959, 0.0...
## $ fractal_dimension_se            <dbl> 0.14597239, 0.06136423, 0.10883324, 0.0...
## $ radius_worst                     <dbl> 0.54108858, 0.19281395, 0.63322661, 0.0...

```



```

#confirming total number of rows = 569
sum(nrow(test_data_bag_all), nrow(train_data_bag_all))

## [1] 569

#create a formula with the variables being analyzed
#set the seed
set.seed(1842)
# Explore all the variables first
form_cancer_bag <- as.formula(diagnosis~radius_mean+texture_mean+perimeter_mean+area_mean+
                                smoothness_mean+compactness_mean+concavity_mean+
                                concave.points_mean+
                                symmetry_mean+fractal_dimension_mean+
                                radius_worst+texture_worst+ perimeter_worst+
                                area_worst+smoothness_worst+
                                compactness_worst+ concavity_worst+concave.points_worst+
                                radius_se+texture_se+perimeter_se+area_se+smoothness_se+compactness_se+concavity_se+
                                concave.points_se+symmetry_se+fractal_dimension_se,
                                data=cancer_bag_all)

#bagged set of trees with mtry=5 (rough square root of predictors), ntree=300
rf1all <- randomForest(form_cancer_bag, data=train_data_bag_all, mtry=5, ntree=300,
                        na.action = na.roughfix)
rf1all

##
## Call:
##  randomForest(formula = form_cancer_bag, data = train_data_bag_all, mtry = 5, ntree = 300, na.action = na.roughfix)
##              Type of random forest: classification
##                      Number of trees: 300
##  No. of variables tried at each split: 5
##
##              OOB estimate of  error rate: 5.05%
##  Confusion matrix:
##    B   M class.error
##  B 282 10 0.03424658
##  M 13 150 0.07975460

#table of importance
importance(rf1all)

##                                     MeanDecreaseGini
## radius_mean                           10.9260923
## texture_mean                          3.0503226
## perimeter_mean                       9.9758531

```

```

## area_mean           7.9869764
## smoothness_mean    1.4983358
## compactness_mean   2.7848510
## concavity_mean     7.9587885
## concave.points_mean 24.4351076
## symmetry_mean      0.9590279
## fractal_dimension_mean 0.5998231
## radius_worst        18.4478417
## texture_worst       3.6264004
## perimeter_worst    29.1628851
## area_worst          24.6521248
## smoothness_worst    2.8533677
## compactness_worst   2.7978557
## concavity_worst     8.2357984
## concave.points_worst 24.7130515
## symmetry_worst      2.2765153
## fractal_dimension_worst 1.4920815
## radius_se            2.8561662
## texture_se            0.7976250
## perimeter_se          3.2792832
## area_se               7.0580392
## smoothness_se          1.0292325
## compactness_se          1.0777282
## concavity_se            1.0280516
## concave.points_se      1.6160159
## symmetry_se             0.9371632
## fractal_dimension_se    0.9977191

test_data_bag_all$rf1allpredict <- predict(rf1all, test_data_bag_all, type="class")

#Construct the confusion matrix for the test data. What percent of the values were misclassified.
table(test_data_bag_all$diagnosis,test_data_bag_all$rf1allpredict)

##
##      B   M
##      B 64   1
##      M   1 48

#table of importance
importance(rf1all)

##                                     MeanDecreaseGini
## radius_mean                  10.9260923
## texture_mean                  3.0503226
## perimeter_mean                9.9758531
## area_mean                     7.9869764
## smoothness_mean                1.4983358
## compactness_mean               2.7848510
## concavity_mean                 7.9587885

```

```

## concave.points_mean           24.4351076
## symmetry_mean                 0.9590279
## fractal_dimension_mean        0.5998231
## radius_worst                  18.4478417
## texture_worst                 3.6264004
## perimeter_worst               29.1628851
## area_worst                    24.6521248
## smoothness_worst              2.8533677
## compactness_worst              2.7978557
## concavity_worst               8.2357984
## concave.points_worst          24.7130515
## symmetry_worst                2.2765153
## fractal_dimension_worst        1.4920815
## radius_se                      2.8561662
## texture_se                      0.7976250
## perimeter_se                   3.2792832
## area_se                         7.0580392
## smoothness_se                  1.0292325
## compactness_se                 1.0777282
## concavity_se                   1.0280516
## concave.points_se              1.6160159
## symmetry_se                     0.9371632
## fractal_dimension_se            0.9977191

rf2all <- randomForest(form_cancer_bag, data=train_data_bag_all, mtry=29, ntree=300,
                        na.action = na.roughfix)
rf2all

##
## Call:
##   randomForest(formula = form_cancer_bag, data = train_data_bag_all,      m
try = 29, ntree = 300, na.action = na.roughfix)
##           Type of random forest: classification
##                   Number of trees: 300
##   No. of variables tried at each split: 29
##
##           OOB estimate of  error rate: 5.05%
## Confusion matrix:
##   B   M class.error
## B 281  11  0.03767123
## M  12 151  0.07361963

test_data_bag_all$rf2allpredict <- predict(rf2all, test_data_bag_all, type="class")
#Construct the confusion matrix for the test data. What percent of the values were misclassified.
table(test_data_bag_all$diagnosis,test_data_bag_all$rf2allpredict)

##
##   B   M

```

```

##   B 62 3
##   M 1 48






```

```

na.action = na.roughfix)
rf1reduced

##
## Call:
##  randomForest(formula = form_cancer_bag2, data = train_data_bag_all,
mtry = 2, ntree = 300, na.action = na.roughfix)
##          Type of random forest: classification
##                      Number of trees: 300
##  No. of variables tried at each split: 2
##
##          OOB estimate of  error rate: 6.81%
## Confusion matrix:
##      B  M class.error
## B 278 14  0.04794521
## M 17 146  0.10429448

test_data_bag_all$rf1reducedpredict <- predict(rf1reduced, test_data_bag_all,
type="class")
#Construct the confusion matrix for the test data. What percent of the values were misclassified.
table(test_data_bag_all$diagnosis,test_data_bag_all$rf1reducedpredict)

##
##      B  M
## B 62  3
## M 4 45

#table of importance
importance(rf1reduced)

##                                MeanDecreaseGini
## concave.points_mean            33.25374
## area_worst                      41.86140
## radius_worst                     47.49678
## perimeter_worst                  38.58774
## concave.points_worst            47.68697

rf2reduced <- randomForest(form_cancer_bag, data=train_data_bag_all, mtry=3,
ntree=300,
na.action = na.roughfix)
rf2reduced

##
## Call:
##  randomForest(formula = form_cancer_bag, data = train_data_bag_all,      m
try = 3, ntree = 300, na.action = na.roughfix)
##          Type of random forest: classification
##                      Number of trees: 300
##  No. of variables tried at each split: 3
##

```

```

##          OOB estimate of  error rate: 4.84%
## Confusion matrix:
##   B   M class.error
## B 283   9  0.03082192
## M 13 150  0.07975460

test_data_bag_all$rf2reducedpredict <- predict(rf2reduced, test_data_bag_all,
type="class")
#Construct the confusion matrix for the test data. What percent of the values were misclassified.
table(test_data_bag_all$diagnosis,test_data_bag_all$rf2reducedpredict)

##
##          B   M
## B 64   1
## M 1 48

#table of importance
importance(rf2reduced)

##                               MeanDecreaseGini
## radius_mean                  11.8970951
## texture_mean                  3.6872957
## perimeter_mean                 9.1863755
## area_mean                     12.7287220
## smoothness_mean                 2.7633897
## compactness_mean                 3.8125162
## concavity_mean                  12.2691550
## concave.points_mean                 15.4306236
## symmetry_mean                  0.9189566
## fractal_dimension_mean                 1.1196676
## radius_worst                   23.2561551
## texture_worst                   3.6556464
## perimeter_worst                  15.8637612
## area_worst                     20.4949715
## smoothness_worst                  4.0333080
## compactness_worst                  3.6633000
## concavity_worst                  9.4914402
## concave.points_worst                 17.6974340
## symmetry_worst                   2.4496468
## fractal_dimension_worst                 2.1126660
## radius_se                       5.7423175
## texture_se                      1.2244302
## perimeter_se                     5.6030283
## area_se                          9.7497029
## smoothness_se                     0.9919818
## compactness_se                     1.5879505
## concavity_se                      2.1452490
## concave.points_se                  2.2375770
## symmetry_se                      1.0906096
## fractal_dimension_se                 1.2968147

```

```

#6 - Build a classification algorithm using Kth Nearest Neighbors.
#Tune the value of K appropriately.
#Create a copy of the dataset for use in the knn analysis
#All variables except the last column X which contains only NAs which will cause errors in the KNN algorithm
cancer_knn_all <- cancer[,-33]
#make the diagnosis variable binary, malignant = 1, benign = 0
cancer_knn_all$diagnosis <- ifelse(cancer_knn_all$diagnosis=='M',1,0)

#then rescale the numeric observations using the rescale function
cancer_knn_all$radius_mean <- rescale_x(cancer_knn_all$radius_mean)
cancer_knn_all$texture_mean <- rescale_x(cancer_knn_all$texture_mean)
cancer_knn_all$perimeter_mean <- rescale_x(cancer_knn_all$perimeter_mean)
cancer_knn_all$area_mean <- rescale_x(cancer_knn_all$area_mean)
cancer_knn_all$smoothness_mean <- rescale_x(cancer_knn_all$smoothness_mean)
cancer_knn_all$compactness_mean <- rescale_x(cancer_knn_all$compactness_mean)
cancer_knn_all$concavity_mean <- rescale_x(cancer_knn_all$concavity_mean)
cancer_knn_all$concave.points_mean <- rescale_x(cancer_knn_all$concave.points_mean)
cancer_knn_all$symmetry_mean <- rescale_x(cancer_knn_all$symmetry_mean)
cancer_knn_all$fractal_dimension_mean <- rescale_x(cancer_knn_all$fractal_dimension_mean)

cancer_knn_all$radius_worst <- rescale_x(cancer_knn_all$radius_worst)
cancer_knn_all$texture_worst <- rescale_x(cancer_knn_all$texture_worst)
cancer_knn_all$perimeter_worst <- rescale_x(cancer_knn_all$perimeter_worst)
cancer_knn_all$area_worst <- rescale_x(cancer_knn_all$area_worst)
cancer_knn_all$smoothness_worst <- rescale_x(cancer_knn_all$smoothness_worst)
cancer_knn_all$compactness_worst <- rescale_x(cancer_knn_all$compactness_worst)
cancer_knn_all$concavity_worst <- rescale_x(cancer_knn_all$concavity_worst)
cancer_knn_all$concave.points_worst <- rescale_x(cancer_knn_all$concave.points_worst)
cancer_knn_all$symmetry_worst <- rescale_x(cancer_knn_all$symmetry_worst)
cancer_knn_all$fractal_dimension_worst <- rescale_x(cancer_knn_all$fractal_dimension_worst)

cancer_knn_all$radius_se <- rescale_x(cancer_knn_all$radius_se)
cancer_knn_all$texture_se <- rescale_x(cancer_knn_all$texture_se)
cancer_knn_all$perimeter_se <- rescale_x(cancer_knn_all$perimeter_se)
cancer_knn_all$area_se <- rescale_x(cancer_knn_all$area_se)
cancer_knn_all$smoothness_se <- rescale_x(cancer_knn_all$smoothness_se)
cancer_knn_all$compactness_se <- rescale_x(cancer_knn_all$compactness_se)
cancer_knn_all$concavity_se <- rescale_x(cancer_knn_all$concavity_se)
cancer_knn_all$concave.points_se <- rescale_x(cancer_knn_all$concave.points_se)
cancer_knn_all$symmetry_se <- rescale_x(cancer_knn_all$symmetry_se)
cancer_knn_all$fractal_dimension_se <- rescale_x(cancer_knn_all$fractal_dimension_se)

```

```

#Create data subsets for various KNN analysis approaches from rescaled cancer
_knn_all dataframe

#ALL variables with a mean calculation
cancer_knn_mean <- cancer_knn_all[,1:12]

#ALL variables with a worst (maximum) calculation
cancer_knn_worst <- cancer_knn_all[,c(1:2,23:32)]

#ALL variables with a standard error calculation
cancer_knn_se <- cancer_knn_all[,c(1:2,13:22)]

#Variables included in the final decision tree model
cancer_knn_dt_final <- cancer_knn_all %>% select("diagnosis", "area_worst", "c
oncave.points_worst", "texture_mean")

#Variables included in the final random forest model (rf1reduced)
cancer_knn_rf1reduced <- cancer_knn_all %>% select("diagnosis", "area_worst",
"concave.points_worst", "radius_worst", "perimeter_worst", "concave.points_me
an")

#KNN analysis using all variables in the original data set

#set the seed
set.seed(1842)

#save the number of rows in data as n
n_knn_all <- nrow(cancer_knn_all)

#create the test_data dataset of 20% of the data number of rows
test_index_knn_all <- sample.int(n_knn_all, size = round(0.2*n_knn_all))
test_data_knn_all <- cancer_knn_all[test_index_knn_all,]

#create the train_data dataset as the remaining 80%
#number of rows
train_data_knn_all <- cancer_knn_all[-test_index_knn_all,]

#glimpse of both datasets
glimpse(test_data_knn_all)

## Observations: 114
## Variables: 32
## $ id                      <int> 8712766, 91903901, 8811842, 89143602, ...
## $ diagnosis                <dbl> 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, ...
## $ radius_mean               <dbl> 0.4964267, 0.2219225, 0.6067017, 0.351...
## $ texture_mean              <dbl> 0.5062563, 0.3486642, 0.4007440, 0.338...

```

```

## $ perimeter_mean          <dbl> 0.4996890, 0.2171239, 0.5936701, 0.360...
## $ area_mean                <dbl> 0.35677625, 0.11567338, 0.46086957, 0.1...
## $ smoothness_mean          <dbl> 0.4718787, 0.4420872, 0.3719419, 0.315...
## $ compactness_mean          <dbl> 0.43224342, 0.23050733, 0.34114472, 0.1...
## $ concavity_mean            <dbl> 0.50585754, 0.09840675, 0.29803187, 0.1...
## $ concave.points_mean       <dbl> 0.51838966, 0.10720676, 0.43195825, 0.1...
## $ symmetry_mean              <dbl> 0.2414141, 0.4035354, 0.5222222, 0.330...
## $ fractal_dimension_mean    <dbl> 0.28833193, 0.30855097, 0.12320977, 0.1...
## $ radius_se                  <dbl> 0.35361217, 0.03447402, 0.30555857, 0.1...
## $ texture_se                  <dbl> 0.232010255, 0.113662482, 0.182505304, 0.1...
## $ perimeter_se                <dbl> 0.31004099, 0.02996749, 0.26999011, 0.1...
## $ area_se                      <dbl> 0.21572363, 0.01594702, 0.21964595, 0.1...
## $ smoothness_se                <dbl> 0.15164701, 0.12027059, 0.17306319, 0.1...
## $ compactness_se                <dbl> 0.25601586, 0.11279178, 0.22109232, 0.1...
## $ concavity_se                  <dbl> 0.11727273, 0.04646465, 0.08702020, 0.1...
## $ concave.points_se            <dbl> 0.29721538, 0.10035992, 0.32430385, 0.1...
## $ symmetry_se                  <dbl> 0.05020544, 0.09298137, 0.15601959, 0.1...
## $ fractal_dimension_se          <dbl> 0.14597239, 0.06136423, 0.10883324, 0.1...
## $ radius_worst                 <dbl> 0.54108858, 0.19281395, 0.63322661, 0.1...
## $ texture_worst                 <dbl> 0.54131113, 0.4474947, 0.4429638, 0.269...
## $ perimeter_worst                <dbl> 0.52238657, 0.18223019, 0.59709149, 0.1...
## $ area_worst                      <dbl> 0.36246559, 0.08980535, 0.44824027, 0.1...
## $ smoothness_worst                <dbl> 0.4386845, 0.5535891, 0.4234960, 0.189...
## $ compactness_worst                <dbl> 0.34511162, 0.26109187, 0.28738442, 0.1...
## $ concavity_worst                 <dbl> 0.39057508, 0.22028754, 0.28889776, 0.1...
## $ concave.points_worst            <dbl> 0.5914089, 0.2790378, 0.6254296, 0.350...
## $ symmetry_worst                  <dbl> 0.11728760, 0.32347723, 0.29666864, 0.1...
## $ fractal_dimension_worst          <dbl> 0.24898334, 0.22602650, 0.18044077, 0.1...

glimpse(train_data_knn_all)

## Observations: 455
## Variables: 32

## $ id                           <int> 842302, 84300903, 84348301, 843786, 84...
## $ diagnosis                     <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ radius_mean                   <dbl> 0.5210374, 0.6014956, 0.2100904, 0.258...
## $ texture_mean                  <dbl> 0.0226581, 0.3902604, 0.3608387, 0.202...
## $ perimeter_mean                 <dbl> 0.5459885, 0.5957432, 0.2335015, 0.267...
## $ area_mean                      <dbl> 0.36373277, 0.44941676, 0.10290562, 0.1...
## $ smoothness_mean                 <dbl> 0.5937528, 0.5143089, 0.8113208, 0.678...
## $ compactness_mean                 <dbl> 0.7920373, 0.4310165, 0.8113613, 0.461...
## $ concavity_mean                  <dbl> 0.70313964, 0.46251172, 0.56560450, 0.1...
## $ concave.points_mean            <dbl> 0.7311133, 0.6356859, 0.5228628, 0.402...
## $ symmetry_mean                   <dbl> 0.6863636, 0.5095960, 0.7762626, 0.518...
## $ fractal_dimension_mean          <dbl> 0.60551811, 0.21124684, 1.00000000, 0.1...
## $ radius_se                      <dbl> 0.35614702, 0.22962158, 0.13909107, 0.1...
## $ texture_se                      <dbl> 0.12046941, 0.09430251, 0.17587518, 0.1...
## $ perimeter_se                   <dbl> 0.36903360, 0.18037035, 0.12665504, 0.1...
## $ area_se                          <dbl> 0.27381126, 0.16292179, 0.03815479, 0.1...
## $ smoothness_se                   <dbl> 0.15929565, 0.15083115, 0.25145324, 0.1...

```

```

## $ compactness_se          <dbl> 0.35139844, 0.28395470, 0.54321507, 0....
## $ concavity_se            <dbl> 0.13568182, 0.09676768, 0.14295455, 0....
## $ concave.points_se       <dbl> 0.3006251, 0.3898466, 0.3536655, 0.215...
## $ symmetry_se              <dbl> 0.31164518, 0.20569032, 0.72814769, 0....
## $ fractal_dimension_se    <dbl> 0.18304244, 0.12700551, 0.28720479, 0....
## $ radius_worst             <dbl> 0.62077552, 0.55638563, 0.24831021, 0....
## $ texture_worst            <dbl> 0.14152452, 0.36007463, 0.38592751, 0....
## $ perimeter_worst          <dbl> 0.66831017, 0.50844166, 0.24134668, 0....
## $ area_worst                <dbl> 0.45069799, 0.37450845, 0.09400806, 0....
## $ smoothness_worst          <dbl> 0.6011358, 0.4835898, 0.9154725, 0.712...
## $ compactness_worst         <dbl> 0.61929156, 0.38537513, 0.81401170, 0....
## $ concavity_worst           <dbl> 0.56861022, 0.35974441, 0.54864217, 0....
## $ concave.points_worst     <dbl> 0.9120275, 0.8350515, 0.8848797, 0.598...
## $ symmetry_worst            <dbl> 0.5984624, 0.4037059, 1.0000000, 0.477...
## $ fractal_dimension_worst   <dbl> 0.41886396, 0.21343303, 0.77371114, 0.....

#confirming total number of rows = 569
sum(nrow(test_data_knn_all), nrow(train_data_knn_all))

## [1] 569

#KNN classification of diagnosis variable & confusion matrices for varying values of k

#k = 1
cancer_knn1_all <- knn(train = train_data_knn_all[-2], test = test_data_knn_all[-2],
                         cl=train_data_knn_all$diagnosis, k = 1)
cancer_knn1_all_cm <- table(cancer_knn1_all,test_data_knn_all$diagnosis)
cancer_knn1_all_cm

##
## cancer_knn1_all  0  1
##                  0 48 26
##                  1 17 23

misclassified_knn1_all <- sum(cancer_knn1_all_cm[1,2], cancer_knn1_all_cm[2,1])

#k = 3
cancer_knn3_all <- knn(train = train_data_knn_all[-2], test = test_data_knn_all[-2],
                         cl=train_data_knn_all$diagnosis, k = 3)
cancer_knn3_all_cm <- table(cancer_knn3_all,test_data_knn_all$diagnosis)
cancer_knn3_all_cm

##
## cancer_knn3_all  0  1
##                  0 54 33
##                  1 11 16

```

```

misclassified_knn3_all <- sum(cancer_knn3_all_cm[1,2], cancer_knn3_all_cm[2,1])

#k approx sqrt(n) = 11
cancer_knn11_all <- knn(train = train_data_knn_all[-2], test = test_data_knn_all[-2],
                           cl=train_data_knn_all$diagnosis, k = 11)
cancer_knn11_all_cm <- table(cancer_knn11_all,test_data_knn_all$diagnosis)
cancer_knn11_all_cm

##
## cancer_knn11_all 0 1
##                 0 61 32
##                 1 4 17

misclassified_knn11_all <- sum(cancer_knn11_all_cm[1,2], cancer_knn11_all_cm[2,1])

#k = 51
cancer_knn51_all <- knn(train = train_data_knn_all[-2], test = test_data_knn_all[-2],
                           cl=train_data_knn_all$diagnosis, k = 11)
cancer_knn51_all_cm <- table(cancer_knn51_all,test_data_knn_all$diagnosis)
cancer_knn51_all_cm

##
## cancer_knn51_all 0 1
##                 0 61 32
##                 1 4 17

misclassified_knn51_all <- sum(cancer_knn51_all_cm[1,2], cancer_knn51_all_cm[2,1])

#k = 101
cancer_knn101_all <- knn(train = train_data_knn_all[-2], test = test_data_knn_all[-2],
                           cl=train_data_knn_all$diagnosis, k = 11)
cancer_knn101_all_cm <- table(cancer_knn101_all,test_data_knn_all$diagnosis)
cancer_knn101_all_cm

##
## cancer_knn101_all 0 1
##                 0 61 32
##                 1 4 17

misclassified_knn101_all <- sum(cancer_knn101_all_cm[1,2], cancer_knn101_all_cm[2,1])

#calculate minimum misclassified count
min_misclassification_knn_all <- min(misclassified_knn1_all, misclassified_knn3_all, misclassified_knn11_all, misclassified_knn51_all, misclassified_knn101_all)

```

```

1_all)

#KNN analysis using only mean calculation variables

#set the seed
set.seed(1842)

#save the number of rows in data as n
n_knn_mean <- nrow(cancer_knn_mean)

#create the test_data dataset of 20% of the data number of rows
test_index_knn_mean <- sample.int(n_knn_mean, size = round(0.2*n_knn_mean))
test_data_knn_mean <- cancer_knn_mean[test_index_knn_mean,]

#create the train_data dataset as the remaining 80%
#number of rows
train_data_knn_mean <- cancer_knn_mean[-test_index_knn_mean,]

#glimpse of both datasets
glimpse(test_data_knn_mean)

## Observations: 114
## Variables: 12
## $ id                  <int> 8712766, 91903901, 8811842, 89143602, 8...
## $ diagnosis           <dbl> 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, ...
## $ radius_mean          <dbl> 0.4964267, 0.2219225, 0.6067017, 0.3516...
## $ texture_mean          <dbl> 0.5062563, 0.3486642, 0.4007440, 0.3388...
## $ perimeter_mean        <dbl> 0.4996890, 0.2171239, 0.5936701, 0.3609...
## $ area_mean              <dbl> 0.35677625, 0.11567338, 0.46086957, 0.2...
## $ smoothness_mean        <dbl> 0.4718787, 0.4420872, 0.3719419, 0.3154...
## $ compactness_mean       <dbl> 0.43224342, 0.23050733, 0.34114472, 0.4...
## $ concavity_mean         <dbl> 0.50585754, 0.09840675, 0.29803187, 0.3...
## $ concave.points_mean   <dbl> 0.51838966, 0.10720676, 0.43195825, 0.3...
## $ symmetry_mean          <dbl> 0.2414141, 0.4035354, 0.5222222, 0.3303...
## $ fractal_dimension_mean <dbl> 0.28833193, 0.30855097, 0.12320977, 0.4...

glimpse(train_data_knn_mean)

## Observations: 455
## Variables: 12
## $ id                  <int> 842302, 84300903, 84348301, 843786, 844...
## $ diagnosis           <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ radius_mean          <dbl> 0.5210374, 0.6014956, 0.2100904, 0.2588...
## $ texture_mean          <dbl> 0.0226581, 0.3902604, 0.3608387, 0.2025...
## $ perimeter_mean        <dbl> 0.5459885, 0.5957432, 0.2335015, 0.2679...
## $ area_mean              <dbl> 0.36373277, 0.44941676, 0.10290562, 0.1...
## $ smoothness_mean        <dbl> 0.5937528, 0.5143089, 0.8113208, 0.6786...
## $ compactness_mean       <dbl> 0.7920373, 0.4310165, 0.8113613, 0.4619...
## $ concavity_mean         <dbl> 0.70313964, 0.46251172, 0.56560450, 0.3...

```

```

## $ concave.points_mean      <dbl> 0.7311133, 0.6356859, 0.5228628, 0.4020...
## $ symmetry_mean           <dbl> 0.6863636, 0.5095960, 0.7762626, 0.5186...
## $ fractal_dimension_mean  <dbl> 0.60551811, 0.21124684, 1.00000000, 0.5...

#confirming total number of rows = 569
sum(nrow(test_data_knn_mean), nrow(train_data_knn_mean))

## [1] 569

#KNN classification of diagnosis variable & confusion matrices for varying values of k

#k = 1
cancer_knn1_mean <- knn(train = train_data_knn_mean[-2], test = test_data_knn_mean[-2],
                           cl=train_data_knn_mean$diagnosis, k = 1)
cancer_knn1_mean_cm <- table(cancer_knn1_mean,test_data_knn_mean$diagnosis)
cancer_knn1_mean_cm

##
## cancer_knn1_mean  0  1
##                  0 48 26
##                  1 17 23

misclassified_knn1_mean <- sum(cancer_knn1_mean_cm[1,2], cancer_knn1_mean_cm[2,1])

#k = 3
cancer_knn3_mean <- knn(train = train_data_knn_mean[-2], test = test_data_knn_mean[-2],
                           cl=train_data_knn_mean$diagnosis, k = 3)
cancer_knn3_mean_cm <- table(cancer_knn3_mean,test_data_knn_mean$diagnosis)
cancer_knn3_mean_cm

##
## cancer_knn3_mean  0  1
##                  0 54 33
##                  1 11 16

misclassified_knn3_mean <- sum(cancer_knn3_mean_cm[1,2], cancer_knn3_mean_cm[2,1])

#k approx sqrt(n) = 11
cancer_knn11_mean <- knn(train = train_data_knn_mean[-2], test = test_data_knn_mean[-2],
                           cl=train_data_knn_mean$diagnosis, k = 11)
cancer_knn11_mean_cm <- table(cancer_knn11_mean,test_data_knn_mean$diagnosis)
cancer_knn11_mean_cm

##
## cancer_knn11_mean  0  1

```

```

##          0 61 32
##          1 4 17

misclassified_knn11_mean <- sum(cancer_knn11_mean_cm[1,2], cancer_knn11_mean_cm[2,1])

#k = 51
cancer_knn51_mean <- knn(train = train_data_knn_mean[-2], test = test_data_knn_mean[-2],
                           cl=train_data_knn_mean$diagnosis, k = 11)
cancer_knn51_mean_cm <- table(cancer_knn51_mean,test_data_knn_mean$diagnosis)
cancer_knn51_mean_cm

##
## cancer_knn51_mean  0  1
##                      0 61 32
##                      1 4 17

misclassified_knn51_mean <- sum(cancer_knn51_mean_cm[1,2], cancer_knn51_mean_cm[2,1])

#k = 101
cancer_knn101_mean <- knn(train = train_data_knn_mean[-2], test = test_data_knn_mean[-2],
                           cl=train_data_knn_mean$diagnosis, k = 11)
cancer_knn101_mean_cm <- table(cancer_knn101_mean,test_data_knn_mean$diagnosis)
cancer_knn101_mean_cm

##
## cancer_knn101_mean  0  1
##                      0 61 32
##                      1 4 17

misclassified_knn101_mean <- sum(cancer_knn101_mean_cm[1,2], cancer_knn101_mean_cm[2,1])

#calculate minimum misclassified count
min_classification_knn_mean <- min(misclassified_knn1_mean, misclassified_knn3_mean, misclassified_knn11_mean, misclassified_knn51_mean, misclassified_knn101_mean)

#KNN analysis using only worst calculation variables

#set the seed
set.seed(1842)

#save the number of rows in data as n

```

```

n_knn_worst <- nrow(cancer_knn_worst)

#create the test_data dataset of 20% of the data number of rows
test_index_knn_worst <- sample.int(n_knn_worst, size = round(0.2*n_knn_worst))
)
test_data_knn_worst <- cancer_knn_worst[test_index_knn_worst,]

#create the train_data dataset as the remaining 80%
#number of rows
train_data_knn_worst <- cancer_knn_worst[-test_index_knn_worst,]

#glimpse of both datasets
glimpse(test_data_knn_worst)

## Observations: 114
## Variables: 12
## $ id                      <int> 8712766, 91903901, 8811842, 89143602, ...
## $ diagnosis                <dbl> 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, ...
## $ radius_worst              <dbl> 0.54108858, 0.19281395, 0.63322661, 0. ...
## $ texture_worst              <dbl> 0.5413113, 0.4474947, 0.4429638, 0.269 ...
## $ perimeter_worst            <dbl> 0.52238657, 0.18223019, 0.59709149, 0. ...
## $ area_worst                 <dbl> 0.36246559, 0.08980535, 0.44824027, 0. ...
## $ smoothness_worst            <dbl> 0.4386845, 0.5535891, 0.4234960, 0.189 ...
## $ compactness_worst           <dbl> 0.34511162, 0.26109187, 0.28738442, 0. ...
## $ concavity_worst              <dbl> 0.39057508, 0.22028754, 0.28889776, 0. ...
## $ concave.points_worst        <dbl> 0.5914089, 0.2790378, 0.6254296, 0.350 ...
## $ symmetry_worst              <dbl> 0.11728760, 0.32347723, 0.29666864, 0. ...
## $ fractal_dimension_worst     <dbl> 0.24898334, 0.22602650, 0.18044077, 0. ...

glimpse(train_data_knn_worst)

## Observations: 455
## Variables: 12
## $ id                      <int> 842302, 84300903, 84348301, 843786, 84...
## $ diagnosis                <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ radius_worst              <dbl> 0.62077552, 0.55638563, 0.24831021, 0. ...
## $ texture_worst              <dbl> 0.14152452, 0.36007463, 0.38592751, 0. ...
## $ perimeter_worst            <dbl> 0.66831017, 0.50844166, 0.24134668, 0. ...
## $ area_worst                 <dbl> 0.45069799, 0.37450845, 0.09400806, 0. ...
## $ smoothness_worst            <dbl> 0.6011358, 0.4835898, 0.9154725, 0.712 ...
## $ compactness_worst           <dbl> 0.61929156, 0.38537513, 0.81401170, 0. ...
## $ concavity_worst              <dbl> 0.56861022, 0.35974441, 0.54864217, 0. ...
## $ concave.points_worst        <dbl> 0.9120275, 0.8350515, 0.8848797, 0.598 ...
## $ symmetry_worst              <dbl> 0.5984624, 0.4037059, 1.0000000, 0.477 ...
## $ fractal_dimension_worst     <dbl> 0.41886396, 0.21343303, 0.77371114, 0. ...

#confirming total number of rows = 569
sum(nrow(test_data_knn_worst), nrow(train_data_knn_worst))

## [1] 569

```

```

#KNN classification of diagnosis variable & confusion matrices for varying values of k

#k = 1
cancer_knn1_worst <- knn(train = train_data_knn_worst[-2], test = test_data_knn_worst[-2],
                           cl=train_data_knn_worst$diagnosis, k = 1)
cancer_knn1_worst_cm <- table(cancer_knn1_worst,test_data_knn_worst$diagnosis)
cancer_knn1_worst_cm

##
## cancer_knn1_worst  0  1
##                      0 48 26
##                      1 17 23

misclassified_knn1_worst <- sum(cancer_knn1_worst_cm[1,2], cancer_knn1_worst_cm[2,1])

#k = 3
cancer_knn3_worst <- knn(train = train_data_knn_worst[-2], test = test_data_knn_worst[-2],
                           cl=train_data_knn_worst$diagnosis, k = 3)
cancer_knn3_worst_cm <- table(cancer_knn3_worst,test_data_knn_worst$diagnosis)
cancer_knn3_worst_cm

##
## cancer_knn3_worst  0  1
##                      0 54 33
##                      1 11 16

misclassified_knn3_worst <- sum(cancer_knn3_worst_cm[1,2], cancer_knn3_worst_cm[2,1])

#k approx sqrt(n) = 11
cancer_knn11_worst <- knn(train = train_data_knn_worst[-2], test = test_data_knn_worst[-2],
                           cl=train_data_knn_worst$diagnosis, k = 11)
cancer_knn11_worst_cm <- table(cancer_knn11_worst,test_data_knn_worst$diagnosis)
cancer_knn11_worst_cm

##
## cancer_knn11_worst  0  1
##                      0 61 32
##                      1  4 17

misclassified_knn11_worst <- sum(cancer_knn11_worst_cm[1,2], cancer_knn11_worst_cm[2,1])

```

```

#k = 51
cancer_knn51_worst <- knn(train = train_data_knn_worst[-2], test = test_data_knn_worst[-2],
                           cl=train_data_knn_worst$diagnosis, k = 11)
cancer_knn51_worst_cm <- table(cancer_knn51_worst,test_data_knn_worst$diagnosis)
cancer_knn51_worst_cm

##
## cancer_knn51_worst  0  1
##                  0 61 32
##                  1  4 17

misclassified_knn51_worst <- sum(cancer_knn51_worst_cm[1,2], cancer_knn51_worst_cm[2,1])

#k = 101
cancer_knn101_worst <- knn(train = train_data_knn_worst[-2], test = test_data_knn_worst[-2],
                           cl=train_data_knn_worst$diagnosis, k = 11)
cancer_knn101_worst_cm <- table(cancer_knn101_worst,test_data_knn_worst$diagnosis)
cancer_knn101_worst_cm

##
## cancer_knn101_worst  0  1
##                  0 61 32
##                  1  4 17

misclassified_knn101_worst <- sum(cancer_knn101_worst_cm[1,2], cancer_knn101_worst_cm[2,1])

#calculate minimum misclassified count
min_misclassification_knn_worst <- min(misclassified_knn1_worst, misclassified_knn3_worst, misclassified_knn11_worst, misclassified_knn51_worst, misclassified_knn101_worst)

#KNN analysis using only standard error calculation variables

#set the seed
set.seed(1842)

#save the number of rows in data as n
n_knn_se <- nrow(cancer_knn_se)

#create the test_data dataset of 20% of the data number of rows
test_index_knn_se <- sample.int(n_knn_se, size = round(0.2*n_knn_se))
test_data_knn_se <- cancer_knn_se[test_index_knn_se,]

```

```

#create the train_data dataset as the remaining 80%
#number of rows
train_data_knn_se <- cancer_knn_se[-test_index_knn_se, ]

#glimpse of both datasets
glimpse(test_data_knn_se)

## Observations: 114
## Variables: 12
## $ id                  <int> 8712766, 91903901, 8811842, 89143602, 867...
## $ diagnosis           <dbl> 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, ...
## $ radius_se            <dbl> 0.35361217, 0.03447402, 0.30555857, 0.278...
## $ texture_se            <dbl> 0.232010255, 0.113662482, 0.182505304, 0.1...
## $ perimeter_se          <dbl> 0.31004099, 0.02996749, 0.26999011, 0.169...
## $ area_se               <dbl> 0.21572363, 0.01594702, 0.21964595, 0.131...
## $ smoothness_se          <dbl> 0.15164701, 0.12027059, 0.17306319, 0.205...
## $ compactness_se          <dbl> 0.25601586, 0.11279178, 0.22109232, 0.782...
## $ concavity_se            <dbl> 0.11727273, 0.04646465, 0.08702020, 0.251...
## $ concave.points_se       <dbl> 0.29721538, 0.10035992, 0.32430385, 0.524...
## $ symmetry_se              <dbl> 0.05020544, 0.09298137, 0.15601959, 0.462...
## $ fractal_dimension_se     <dbl> 0.14597239, 0.06136423, 0.10883324, 0.758...

glimpse(train_data_knn_se)

## Observations: 455
## Variables: 12
## $ id                  <int> 842302, 84300903, 84348301, 843786, 84458...
## $ diagnosis           <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, ...
## $ radius_se            <dbl> 0.35614702, 0.22962158, 0.13909107, 0.080...
## $ texture_se            <dbl> 0.12046941, 0.09430251, 0.17587518, 0.117...
## $ perimeter_se          <dbl> 0.36903360, 0.18037035, 0.12665504, 0.068...
## $ area_se               <dbl> 0.27381126, 0.16292179, 0.03815479, 0.038...
## $ smoothness_se          <dbl> 0.15929565, 0.15083115, 0.25145324, 0.197...
## $ compactness_se          <dbl> 0.35139844, 0.28395470, 0.54321507, 0.234...
## $ concavity_se            <dbl> 0.13568182, 0.09676768, 0.14295455, 0.092...
## $ concave.points_se       <dbl> 0.3006251, 0.3898466, 0.3536655, 0.215381...
## $ symmetry_se              <dbl> 0.31164518, 0.20569032, 0.72814769, 0.193...
## $ fractal_dimension_se     <dbl> 0.18304244, 0.12700551, 0.28720479, 0.144...

#confirming total number of rows = 569
sum(nrow(test_data_knn_se), nrow(train_data_knn_se))

## [1] 569

#KNN classification of diagnosis variable & confusion matrices for varying values of k

#k = 1
cancer_knn1_se <- knn(train = train_data_knn_se[-2], test = test_data_knn_se[-2],

```

```

            cl=train_data_knn_se$diagnosis, k = 1)
cancer_knn1_se_cm <- table(cancer_knn1_se,test_data_knn_se$diagnosis)
cancer_knn1_se_cm

##
## cancer_knn1_se  0  1
##                 0 48 26
##                 1 17 23

misclassified_knn1_se <- sum(cancer_knn1_se_cm[1,2], cancer_knn1_se_cm[2,1])

#k = 3
cancer_knn3_se <- knn(train = train_data_knn_se[-2], test = test_data_knn_se[-2],
                        cl=train_data_knn_se$diagnosis, k = 3)
cancer_knn3_se_cm <- table(cancer_knn3_se,test_data_knn_se$diagnosis)
cancer_knn3_se_cm

##
## cancer_knn3_se  0  1
##                 0 54 33
##                 1 11 16

misclassified_knn3_se <- sum(cancer_knn3_se_cm[1,2], cancer_knn3_se_cm[2,1])

#k approx sqrt(n) = 11
cancer_knn11_se <- knn(train = train_data_knn_se[-2], test = test_data_knn_se[-2],
                        cl=train_data_knn_se$diagnosis, k = 11)
cancer_knn11_se_cm <- table(cancer_knn11_se,test_data_knn_se$diagnosis)
cancer_knn11_se_cm

##
## cancer_knn11_se  0  1
##                 0 61 32
##                 1  4 17

misclassified_knn11_se <- sum(cancer_knn11_se_cm[1,2], cancer_knn11_se_cm[2,1])

#k = 51
cancer_knn51_se <- knn(train = train_data_knn_se[-2], test = test_data_knn_se[-2],
                        cl=train_data_knn_se$diagnosis, k = 11)
cancer_knn51_se_cm <- table(cancer_knn51_se,test_data_knn_se$diagnosis)
cancer_knn51_se_cm

##
## cancer_knn51_se  0  1
##                 0 61 32
##                 1  4 17

```

```

misclassified_knn51_se <- sum(cancer_knn51_se_cm[1,2], cancer_knn51_se_cm[2,1])
])

#k = 101
cancer_knn101_se <- knn(train = train_data_knn_se[-2], test = test_data_knn_se[-2],
                           cl=train_data_knn_se$diagnosis, k = 11)
cancer_knn101_se_cm <- table(cancer_knn101_se, test_data_knn_se$diagnosis)
cancer_knn101_se_cm

##
## cancer_knn101_se  0  1
##                  0 61 32
##                  1  4 17

misclassified_knn101_se <- sum(cancer_knn101_se_cm[1,2], cancer_knn101_se_cm[2,1])

#calculate minimum misclassified count
min_classification_knn_se <- min(misclassified_knn1_se, misclassified_knn3_se, misclassified_knn11_se, misclassified_knn51_se, misclassified_knn101_se)

#KNN analysis using only the variables identified in the final decision tree model

#set the seed
set.seed(1842)

#save the number of rows in data as n
n_knn_dt_final <- nrow(cancer_knn_dt_final)

#create the test_data dataset of 20% of the data number of rows
test_index_knn_dt_final <- sample.int(n_knn_dt_final, size = round(0.2*n_knn_dt_final))
test_data_knn_dt_final <- cancer_knn_dt_final[test_index_knn_dt_final,]

#create the train_data dataset as the remaining 80%
#number of rows
train_data_knn_dt_final <- cancer_knn_dt_final[-test_index_knn_dt_final,]

#glimpse of both datasets
glimpse(test_data_knn_dt_final)

## Observations: 114
## Variables: 4
## $ diagnosis      <dbl> 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, ...
## $ area_worst     <dbl> 0.36246559, 0.08980535, 0.44824027, 0.143...

```



```

#k approx sqrt(n) = 11
cancer_knn11_dt_final <- knn(train = train_data_knn_dt_final[-2], test = test
_data_knn_dt_final[-2],
                                cl=train_data_knn_dt_final$diagnosis, k = 11)
cancer_knn11_dt_final_cm <- table(cancer_knn11_dt_final,test_data_knn_dt_fina
l$diagnosis)
cancer_knn11_dt_final_cm

##
## cancer_knn11_dt_final  0  1
##                      0 65  0
##                      1  0 49

misclassified_knn11_dt_final <- sum(cancer_knn11_dt_final_cm[1,2], cancer_knn
11_dt_final_cm[2,1])

#k = 51
cancer_knn51_dt_final <- knn(train = train_data_knn_dt_final[-2], test = test
_data_knn_dt_final[-2],
                                cl=train_data_knn_dt_final$diagnosis, k = 11)
cancer_knn51_dt_final_cm <- table(cancer_knn51_dt_final,test_data_knn_dt_fina
l$diagnosis)
cancer_knn51_dt_final_cm

##
## cancer_knn51_dt_final  0  1
##                      0 65  0
##                      1  0 49

misclassified_knn51_dt_final <- sum(cancer_knn51_dt_final_cm[1,2], cancer_knn
51_dt_final_cm[2,1])

#k = 101
cancer_knn101_dt_final <- knn(train = train_data_knn_dt_final[-2], test = tes
t_data_knn_dt_final[-2],
                                cl=train_data_knn_dt_final$diagnosis, k = 11)
cancer_knn101_dt_final_cm <- table(cancer_knn101_dt_final,test_data_knn_dt_fi
nal$diagnosis)
cancer_knn101_dt_final_cm

##
## cancer_knn101_dt_final  0  1
##                      0 65  0
##                      1  0 49

misclassified_knn101_dt_final <- sum(cancer_knn101_dt_final_cm[1,2], cancer_k
nn101_dt_final_cm[2,1])

#calculate minimum misclassified count
min_misclassification_knn_dt_final <- min(misclassified_knn1_dt_final, miscla
ssified_knn3_dt_final, misclassified_knn11_dt_final, misclassified_knn51_dt_f

```

```

inal, misclassified_knn101_dt_final)

#KNN analysis using only the variables identified in the final random forrest
model

#set the seed
set.seed(1842)

#save the number of rows in data as n
n_knn_rf1reduced <- nrow(cancer_knn_rf1reduced)

#create the test_data dataset of 20% of the data number of rows
test_index_knn_rf1reduced <- sample.int(n_knn_rf1reduced, size = round(0.2*n_knn_rf1reduced))
test_data_knn_rf1reduced <- cancer_knn_rf1reduced[test_index_knn_rf1reduced,]

#create the train_data dataset as the remaining 80%
#number of rows
train_data_knn_rf1reduced <- cancer_knn_rf1reduced[-test_index_knn_rf1reduced,]

#glimpse of both datasets
glimpse(test_data_knn_rf1reduced)

## Observations: 114
## Variables: 6
## $ diagnosis           <dbl> 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, ...
## $ area_worst          <dbl> 0.36246559, 0.08980535, 0.44824027, 0.143...
## $ concave.points_worst <dbl> 0.5914089, 0.2790378, 0.6254296, 0.350859...
## $ radius_worst         <dbl> 0.54108858, 0.19281395, 0.63322661, 0.278...
## $ perimeter_worst      <dbl> 0.52238657, 0.18223019, 0.59709149, 0.255...
## $ concave.points_mean  <dbl> 0.51838966, 0.10720676, 0.43195825, 0.328...

glimpse(train_data_knn_rf1reduced)

## Observations: 455
## Variables: 6
## $ diagnosis           <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, ...
## $ area_worst          <dbl> 0.45069799, 0.37450845, 0.09400806, 0.136...
## $ concave.points_worst <dbl> 0.9120275, 0.8350515, 0.8848797, 0.598281...
## $ radius_worst         <dbl> 0.62077552, 0.55638563, 0.24831021, 0.268...
## $ perimeter_worst      <dbl> 0.66831017, 0.50844166, 0.24134668, 0.263...
## $ concave.points_mean  <dbl> 0.7311133, 0.6356859, 0.5228628, 0.402037...

#confirming total number of rows = 569
sum(nrow(test_data_knn_rf1reduced), nrow(train_data_knn_rf1reduced))

## [1] 569

```

```

#KNN classification of diagnosis variable & confusion matrices for varying values of k

#k = 1
cancer_knn1_rf1reduced <- knn(train = train_data_knn_rf1reduced[-2], test = test_data_knn_rf1reduced[-2],
                                  cl=train_data_knn_rf1reduced$diagnosis, k = 1)
cancer_knn1_rf1reduced_cm <- table(cancer_knn1_rf1reduced, test_data_knn_rf1reduced$diagnosis)
cancer_knn1_rf1reduced_cm

##
## cancer_knn1_rf1reduced  0  1
##                      0 65  0
##                      1  0 49

misclassified_knn1_rf1reduced <- sum(cancer_knn1_rf1reduced_cm[1,2], cancer_knn1_rf1reduced_cm[2,1])

#k = 3
cancer_knn3_rf1reduced <- knn(train = train_data_knn_rf1reduced[-2], test = test_data_knn_rf1reduced[-2],
                                  cl=train_data_knn_rf1reduced$diagnosis, k = 3)
cancer_knn3_rf1reduced_cm <- table(cancer_knn3_rf1reduced, test_data_knn_rf1reduced$diagnosis)
cancer_knn3_rf1reduced_cm

##
## cancer_knn3_rf1reduced  0  1
##                      0 65  0
##                      1  0 49

misclassified_knn3_rf1reduced <- sum(cancer_knn3_rf1reduced_cm[1,2], cancer_knn3_rf1reduced_cm[2,1])

#k approx sqrt(n) = 11
cancer_knn11_rf1reduced <- knn(train = train_data_knn_rf1reduced[-2], test = test_data_knn_rf1reduced[-2],
                                  cl=train_data_knn_rf1reduced$diagnosis, k = 11)
cancer_knn11_rf1reduced_cm <- table(cancer_knn11_rf1reduced, test_data_knn_rf1reduced$diagnosis)
cancer_knn11_rf1reduced_cm

##
## cancer_knn11_rf1reduced  0  1
##                      0 65  0
##                      1  0 49

misclassified_knn11_rf1reduced <- sum(cancer_knn11_rf1reduced_cm[1,2], cancer_knn11_rf1reduced_cm[2,1])

```

```

#k = 51
cancer_knn51_rf1reduced <- knn(train = train_data_knn_rf1reduced[-2], test =
test_data_knn_rf1reduced[-2],
                                   cl=train_data_knn_rf1reduced$diagnosis, k = 11)
cancer_knn51_rf1reduced_cm <- table(cancer_knn51_rf1reduced,test_data_knn_rf1
reduced$diagnosis)
cancer_knn51_rf1reduced_cm

##
## cancer_knn51_rf1reduced  0  1
##                      0 65  0
##                      1  0 49

misclassified_knn51_rf1reduced <- sum(cancer_knn51_rf1reduced_cm[1,2], cancer
_knn51_rf1reduced_cm[2,1])

#k = 101
cancer_knn101_rf1reduced <- knn(train = train_data_knn_rf1reduced[-2], test =
test_data_knn_rf1reduced[-2],
                                   cl=train_data_knn_rf1reduced$diagnosis, k = 11)
cancer_knn101_rf1reduced_cm <- table(cancer_knn101_rf1reduced,test_data_knn_r
f1reduced$diagnosis)
cancer_knn101_rf1reduced_cm

##
## cancer_knn101_rf1reduced  0  1
##                      0 65  0
##                      1  0 49

misclassified_knn101_rf1reduced <- sum(cancer_knn101_rf1reduced_cm[1,2], canc
er_knn101_rf1reduced_cm[2,1])

#calculate minimum misclassified count
min_misclassification_knn_rf1reduced <- min(misclassified_knn1_rf1reduced, mi
sclassified_knn3_rf1reduced, misclassified_knn11_rf1reduced, misclassified_kn
n51_rf1reduced, misclassified_knn101_rf1reduced)

```