EDA LA2

**shruthi s rao 1nt20is158**

2023-05-22

library(readxl)  
telecom <- read\_excel("C:/Users/user/Downloads/telecom.xlsx")  
View(telecom)  
summary(telecom)

## customerID gender SeniorCitizen Partner   
## Length:7043 Length:7043 Min. :0.0000 Length:7043   
## Class :character Class :character 1st Qu.:0.0000 Class :character   
## Mode :character Mode :character Median :0.0000 Mode :character   
## Mean :0.1621   
## 3rd Qu.:0.0000   
## Max. :1.0000   
##   
## Dependents tenure PhoneService MultipleLines   
## Length:7043 Min. : 0.00 Length:7043 Length:7043   
## Class :character 1st Qu.: 9.00 Class :character Class :character   
## Mode :character Median :29.00 Mode :character Mode :character   
## Mean :32.37   
## 3rd Qu.:55.00   
## Max. :72.00   
##   
## InternetService OnlineSecurity OnlineBackup DeviceProtection   
## Length:7043 Length:7043 Length:7043 Length:7043   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## TechSupport StreamingTV StreamingMovies Contract   
## Length:7043 Length:7043 Length:7043 Length:7043   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## PaperlessBilling PaymentMethod MonthlyCharges TotalCharges   
## Length:7043 Length:7043 Min. : 18.25 Min. : 18.8   
## Class :character Class :character 1st Qu.: 35.50 1st Qu.: 401.4   
## Mode :character Mode :character Median : 70.35 Median :1397.5   
## Mean : 64.76 Mean :2283.3   
## 3rd Qu.: 89.85 3rd Qu.:3794.7   
## Max. :118.75 Max. :8684.8   
## NA's :11   
## Churn   
## Length:7043   
## Class :character   
## Mode :character   
##   
##   
##   
##

str(telecom)

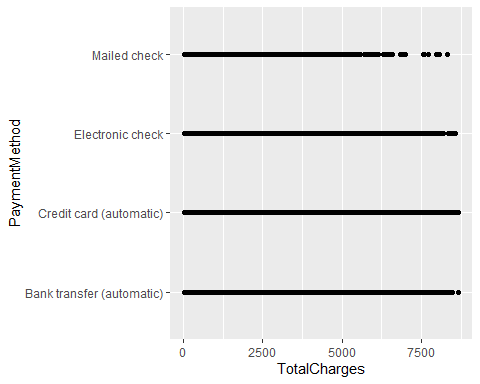
## tibble [7,043 × 21] (S3: tbl\_df/tbl/data.frame)  
## $ customerID : chr [1:7043] "7590-VHVEG" "5575-GNVDE" "3668-QPYBK" "7795-CFOCW" ...  
## $ gender : chr [1:7043] "Female" "Male" "Male" "Male" ...  
## $ SeniorCitizen : num [1:7043] 0 0 0 0 0 0 0 0 0 0 ...  
## $ Partner : chr [1:7043] "Yes" "No" "No" "No" ...  
## $ Dependents : chr [1:7043] "No" "No" "No" "No" ...  
## $ tenure : num [1:7043] 1 34 2 45 2 8 22 10 28 62 ...  
## $ PhoneService : chr [1:7043] "No" "Yes" "Yes" "No" ...  
## $ MultipleLines : chr [1:7043] "No phone service" "No" "No" "No phone service" ...  
## $ InternetService : chr [1:7043] "DSL" "DSL" "DSL" "DSL" ...  
## $ OnlineSecurity : chr [1:7043] "No" "Yes" "Yes" "Yes" ...  
## $ OnlineBackup : chr [1:7043] "Yes" "No" "Yes" "No" ...  
## $ DeviceProtection: chr [1:7043] "No" "Yes" "No" "Yes" ...  
## $ TechSupport : chr [1:7043] "No" "No" "No" "Yes" ...  
## $ StreamingTV : chr [1:7043] "No" "No" "No" "No" ...  
## $ StreamingMovies : chr [1:7043] "No" "No" "No" "No" ...  
## $ Contract : chr [1:7043] "Month-to-month" "One year" "Month-to-month" "One year" ...  
## $ PaperlessBilling: chr [1:7043] "Yes" "No" "Yes" "No" ...  
## $ PaymentMethod : chr [1:7043] "Electronic check" "Mailed check" "Mailed check" "Bank transfer (automatic)" ...  
## $ MonthlyCharges : num [1:7043] 29.9 57 53.9 42.3 70.7 ...  
## $ TotalCharges : num [1:7043] 29.9 1889.5 108.2 1840.8 151.7 ...  
## $ Churn : chr [1:7043] "No" "No" "Yes" "No" ...

library(ggplot2)  
ggplot(telecom,aes(x=gender,y=MonthlyCharges))+geom\_point()

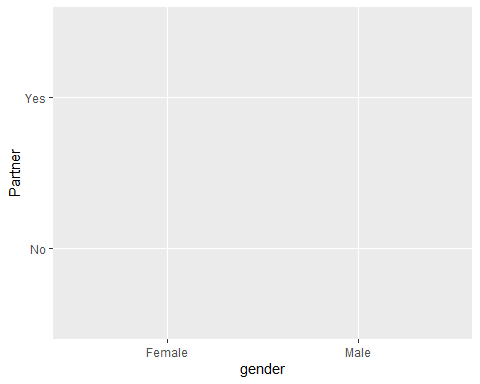


ggplot(telecom,aes(x=TotalCharges,y=PaymentMethod))+geom\_point()

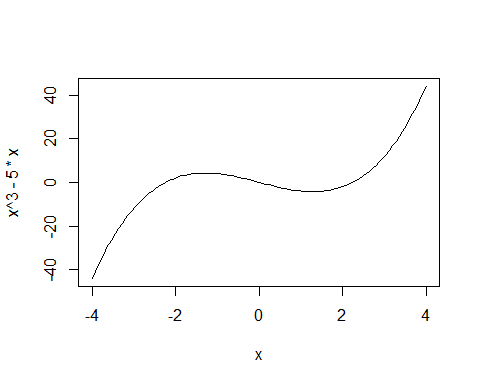
## Warning: Removed 11 rows containing missing values (`geom\_point()`).



library(ggplot2)  
ggplot(telecom,aes(x=gender,y=Partner))+geom\_line()



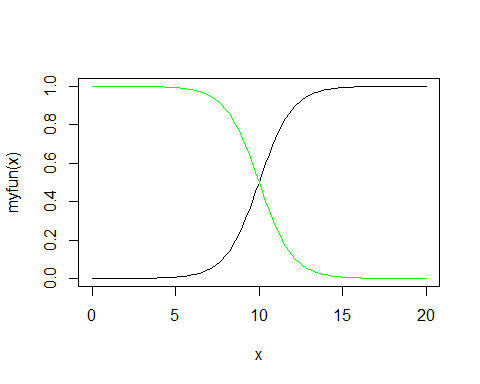
curve(x^3-5\*x,from=-4,to=4)



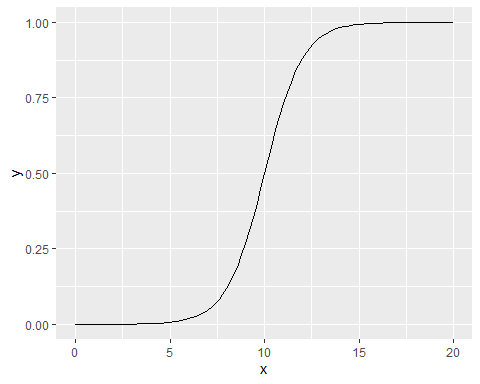
ggplot(telecom,aes(x=gender,y=MonthlyCharges))+geom\_line()+geom\_point()



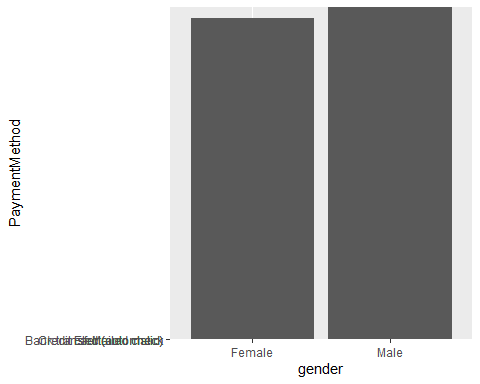
##plot a user defined function  
myfun<-function(xvar){  
 1/( 1+exp(-xvar+10))  
}  
curve(myfun(x),from=0,to=20)  
##add a line  
curve(1 - myfun(x),add=TRUE,col="green")



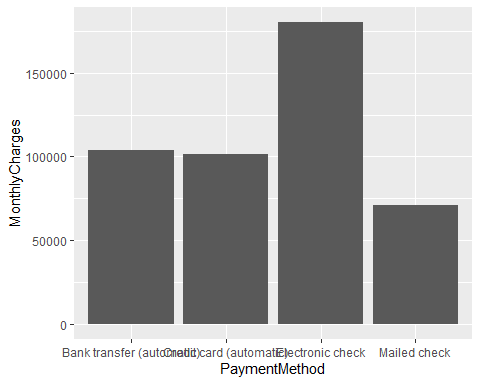
library(ggplot2)  
ggplot(data.frame(x=c(0,20)),aes(x=x))+stat\_function(fun=myfun,geom="line")



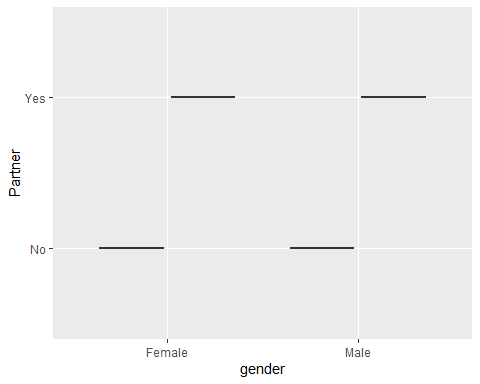
library(gcookbook)  
ggplot(telecom,aes(x=gender,y=PaymentMethod))+geom\_col()



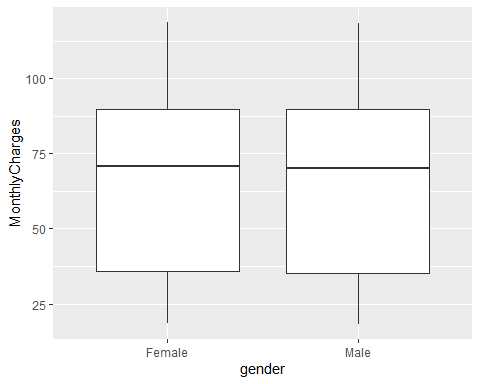
library(gcookbook)  
ggplot(telecom,aes(x=PaymentMethod,y=MonthlyCharges))+geom\_col()



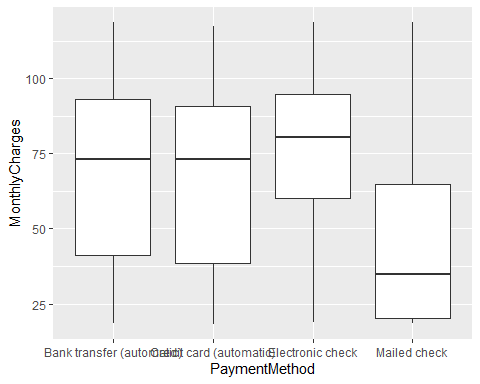
library(ggplot2)  
ggplot(telecom,aes(x=gender,y=Partner))+geom\_boxplot()



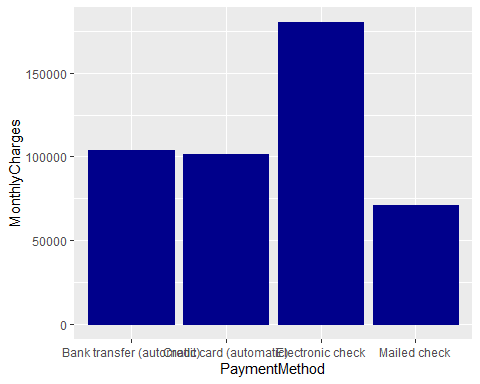
library(ggplot2)  
ggplot(telecom,aes(x=gender,y=MonthlyCharges))+geom\_boxplot()



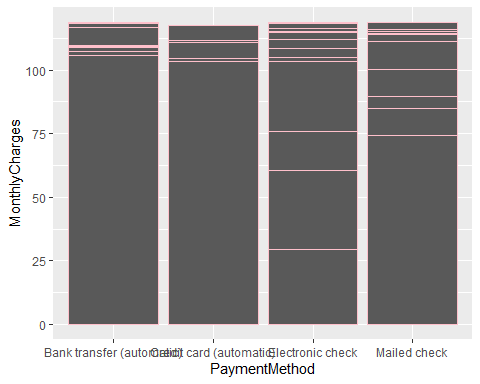
library(ggplot2)  
ggplot(telecom,aes(x=PaymentMethod,y=MonthlyCharges))+geom\_boxplot()



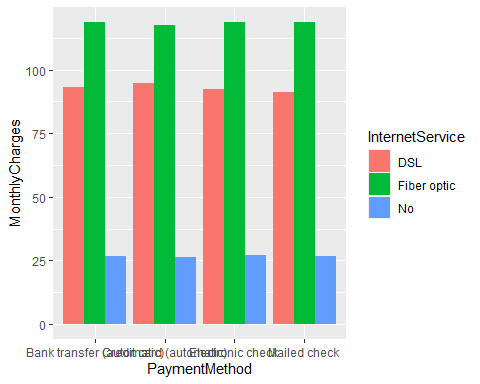
ggplot(telecom,aes(x=PaymentMethod,y=MonthlyCharges))+geom\_col(colour="darkblue")



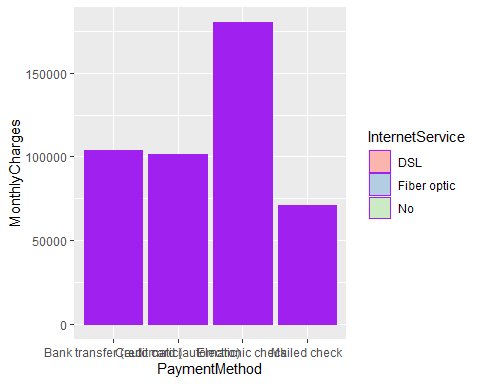
library(gcookbook)  
ggplot(telecom,aes(x=PaymentMethod,y=MonthlyCharges))+geom\_col(position="dodge",colour="pink")



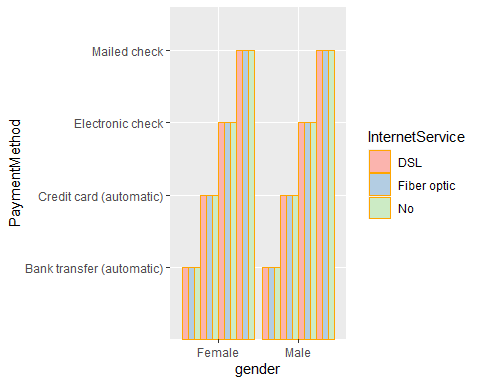
ggplot(telecom,aes(x=PaymentMethod,y=MonthlyCharges, fill=InternetService ))+geom\_col(position="dodge")



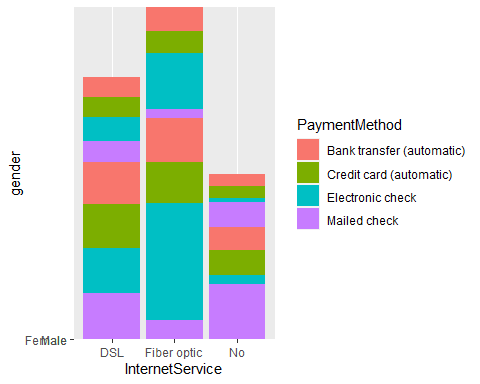
ggplot(telecom,aes(x=PaymentMethod,y=MonthlyCharges,fill=InternetService))+geom\_col(colour="purple")+scale\_fill\_brewer(palette="Pastel1")



ggplot(telecom,aes(x=gender,y=PaymentMethod,fill=InternetService))+geom\_col(position="dodge",colour="orange")+scale\_fill\_brewer(palette="Pastel1")



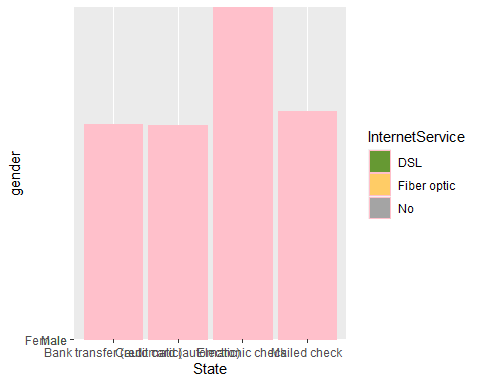
ggplot(telecom,aes(x=InternetService,y=gender,fill=PaymentMethod))+geom\_col()



ggplot(telecom,aes(x=reorder(PaymentMethod,gender),y=gender,fill=InternetService))+geom\_col(colour="pink")+scale\_fill\_manual(values=c('#669933','#FFCC66',"#A4A4A4"))+xlab("State")

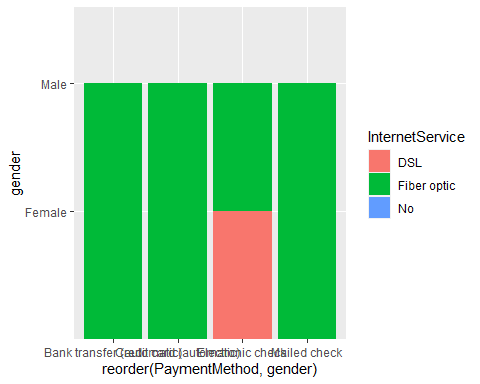
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA

## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA



ggplot(telecom,aes(x=reorder(PaymentMethod,gender),y=gender,fill=InternetService))+geom\_col(position="identity")

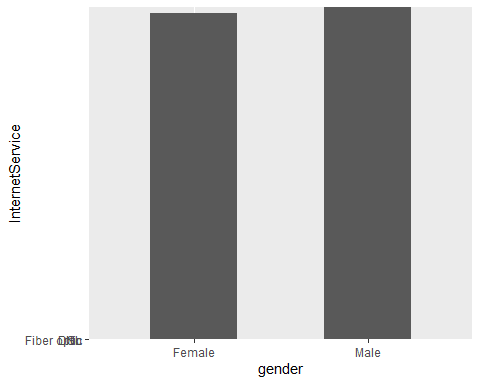
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA



#adjusting bar width and spacing  
library(gcookbook)  
ggplot(telecom,aes(x=gender,y=InternetService))+geom\_col()



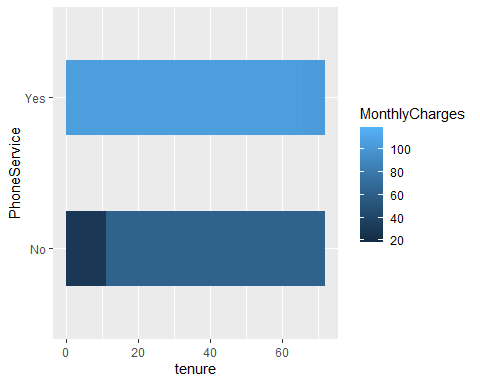
ggplot(telecom,aes(x=gender,y=InternetService))+geom\_col(width=0.5)



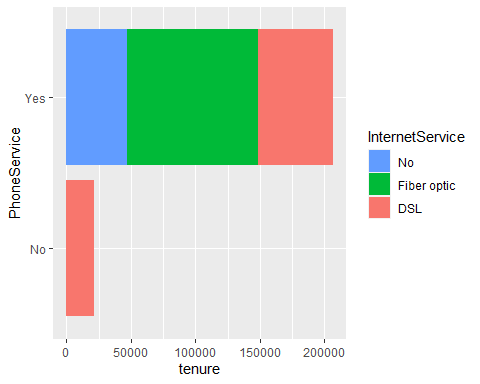
ggplot(telecom,aes(x=gender,y=InternetService))+geom\_col(width=0.8)



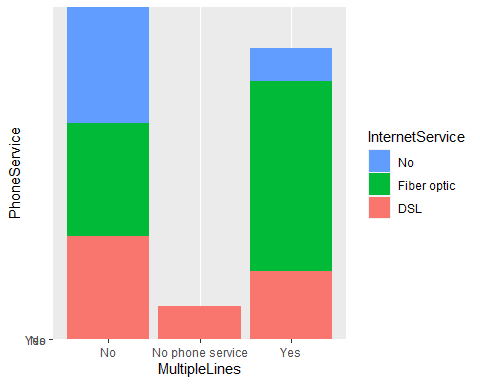
ggplot(telecom,aes(x=tenure,y=PhoneService,fill=MonthlyCharges))+geom\_col(width=0.5,position=position\_dodge(0.5))



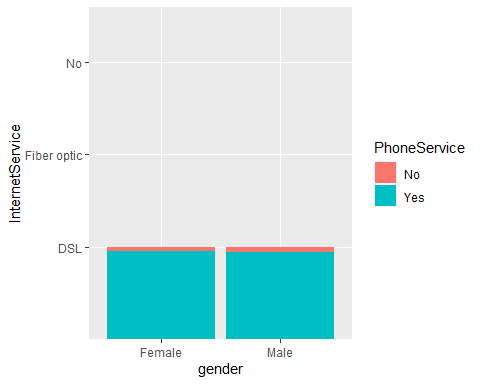
ggplot(telecom,aes(x=tenure,y=PhoneService,fill=InternetService))+geom\_col()+guides(fill=guide\_legend(reverse=TRUE))



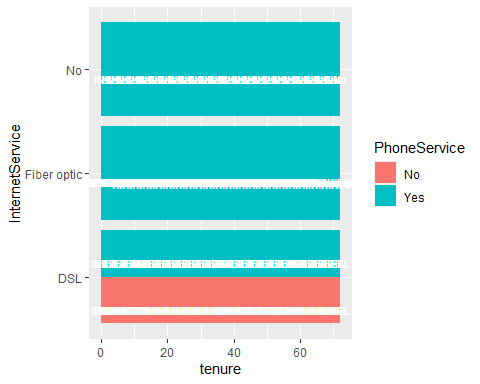
ggplot(telecom,aes(x=MultipleLines,y=PhoneService,fill=InternetService))+geom\_col(position=position\_stack(reverse=TRUE))+guides(fill=guide\_legend(reverse=TRUE))



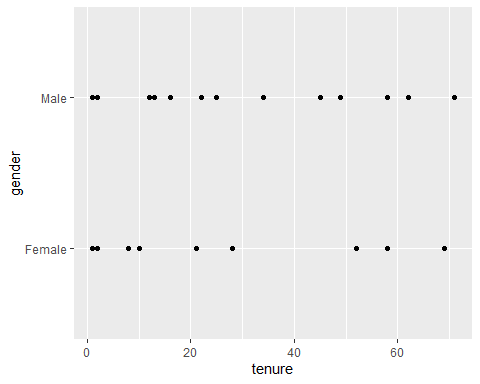
##stacked bar graphs  
library(gcookbook)  
ggplot(telecom,aes(x=gender,y=InternetService,fill=PhoneService))+geom\_col(position="fill")



ggplot(telecom,aes(x=tenure,y=InternetService,fill=PhoneService))+geom\_col(position="dodge")+geom\_text(aes(label=InternetService),colour="white",size=3,vjust=1.5,position=position\_dodge(.9))

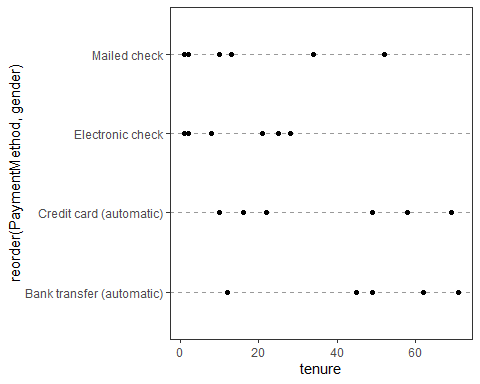


library(gcookbook)  
telecom<-telecom[1:25,]  
ggplot(telecom,aes(x=tenure,y=gender))+geom\_point()



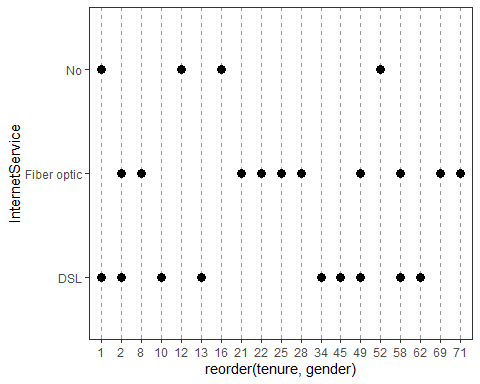
ggplot(telecom,aes(x=tenure,y=reorder(PaymentMethod,gender)))+geom\_point(tenure=3)+theme\_bw()+  
 theme( panel.grid.major.x=element\_blank(),panel.grid.minor.x=element\_blank(),panel.grid.major.y=element\_line(colour="grey60",linetype="dashed"))

## Warning in geom\_point(tenure = 3): Ignoring unknown parameters: `tenure`  
## argument is not numeric or logical: returning NA  
## argument is not numeric or logical: returning NA  
## argument is not numeric or logical: returning NA  
## argument is not numeric or logical: returning NA  
## argument is not numeric or logical: returning NA  
## argument is not numeric or logical: returning NA  
## argument is not numeric or logical: returning NA  
## argument is not numeric or logical: returning NA

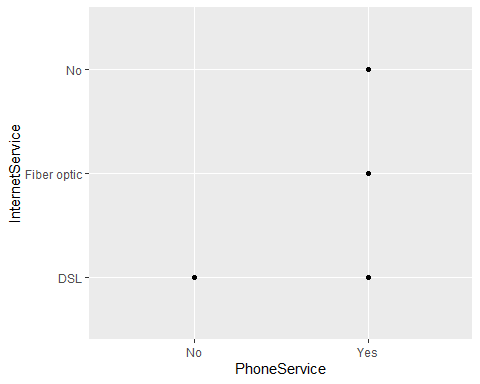


ggplot(telecom,aes(x=reorder(tenure,gender),y=InternetService))+geom\_point(size=3)+theme\_bw()+theme( panel.grid.major.y=element\_blank(),panel.grid.minor.y=element\_blank(),panel.grid.major.x=element\_line(colour="grey60",linetype="dashed"))

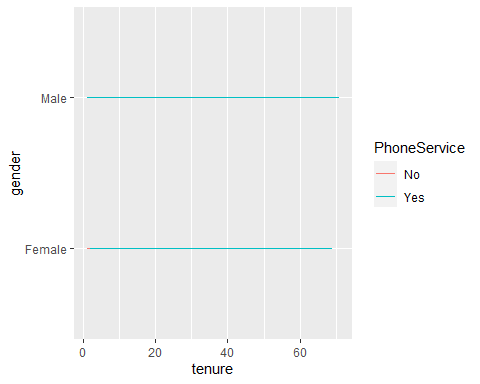
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA  
  
## Warning in mean.default(X[[i]], ...): argument is not numeric or logical:  
## returning NA



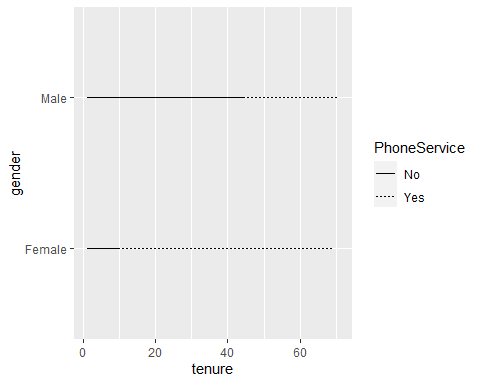
ggplot(telecom,aes(x=PhoneService,y=InternetService))+geom\_line()+geom\_point()



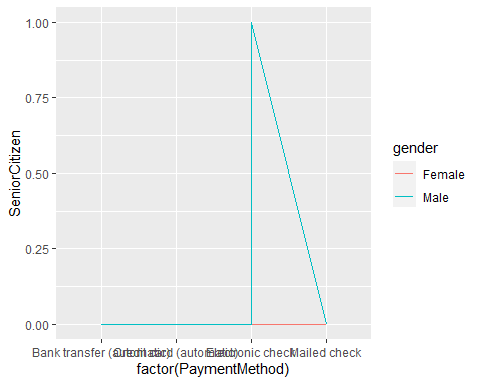
library(gcookbook)  
ggplot(telecom,aes(x=tenure,y=gender,colour=PhoneService))+geom\_line()



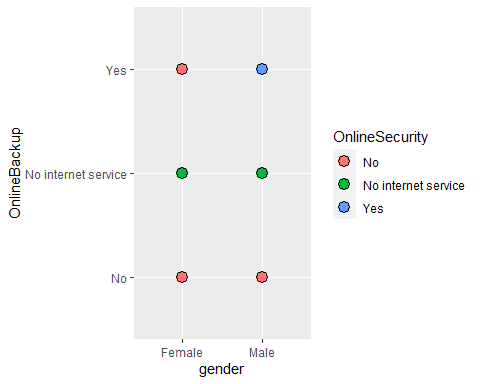
ggplot(telecom,aes(x=tenure,y=gender,linetype=PhoneService))+geom\_line()



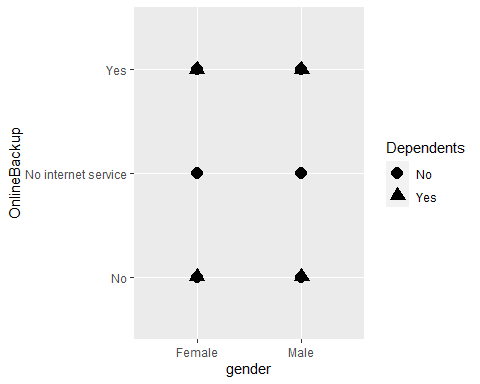
ggplot(telecom,aes(x=factor(PaymentMethod),y=SeniorCitizen,colour=gender,group=gender))+geom\_line()



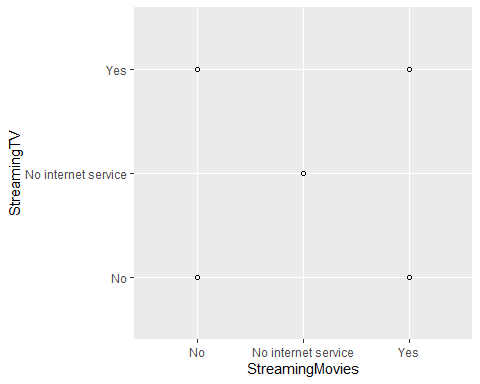
ggplot(telecom,aes(x=gender,y=OnlineBackup,fill=OnlineSecurity))+geom\_line()+geom\_point(size=4,shape=21)



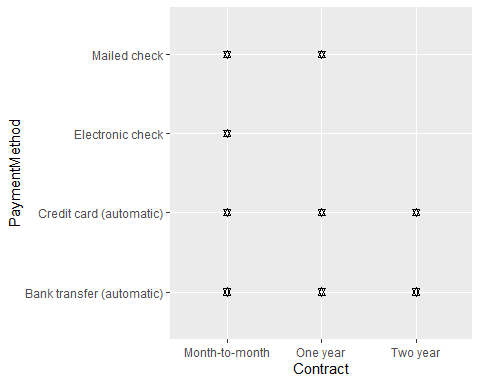
ggplot(telecom,aes(x=gender,y=OnlineBackup,shape=Dependents))+geom\_line()+geom\_point(size=4)



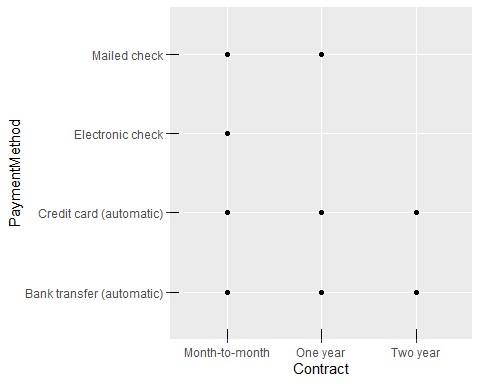
ggplot(telecom,aes(x=StreamingMovies,y=StreamingTV))+geom\_point(shape=21)



ggplot(telecom,aes(x=Contract,y=PaymentMethod))+geom\_point(shape=11)

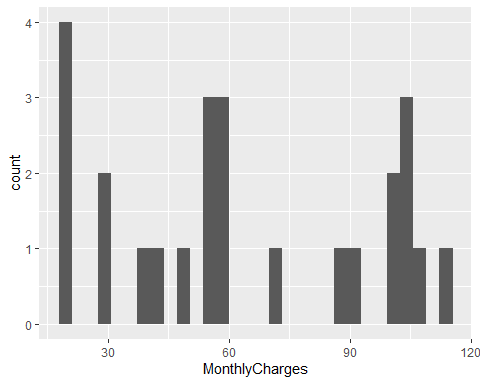


ggplot(telecom,aes(x=Contract,y=PaymentMethod))+geom\_point()+geom\_rug()

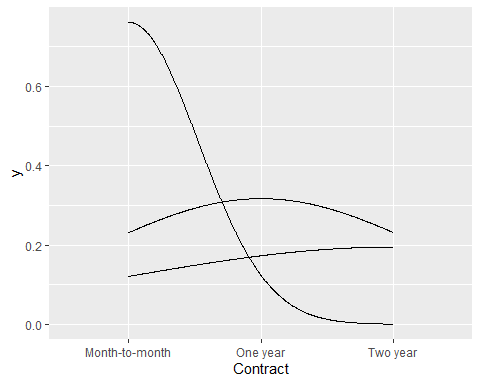


ggplot(telecom,aes(x=MonthlyCharges))+geom\_histogram()

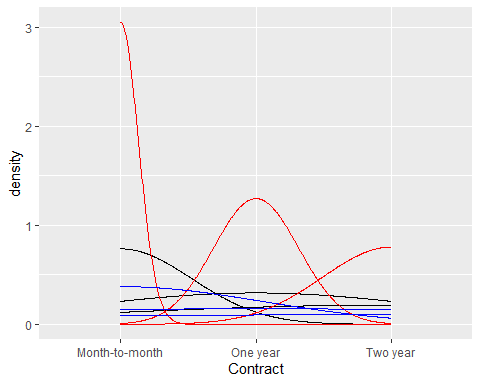
## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



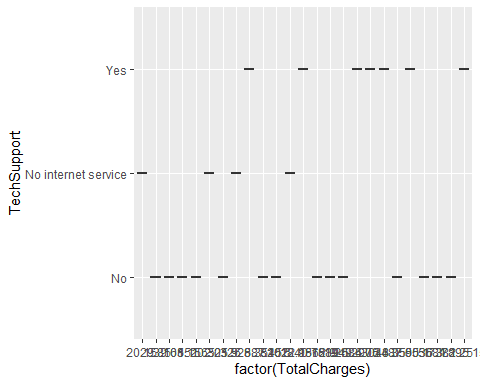
ggplot(telecom,aes(x=Contract))+geom\_line(stat="density")+expand\_limits(y=0)



ggplot(telecom,aes(x=Contract))+geom\_line(stat="density")+geom\_line(stat="density",adjust=.25,colour="red")+geom\_line(stat="density",adjust=2,colour="blue")



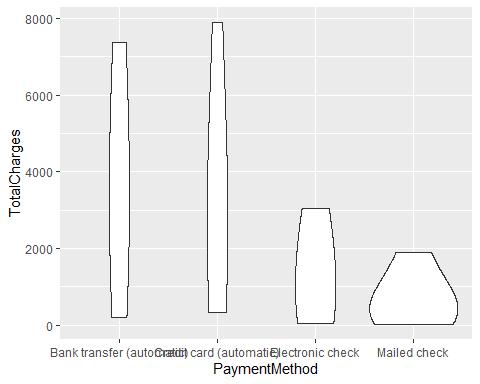
library(MASS)  
ggplot(telecom,aes(x=factor(TotalCharges),y=TechSupport))+geom\_boxplot()



library(gcookbook)  
telecom<-ggplot(telecom,aes(x=PaymentMethod,y=TotalCharges))  
telecom+geom\_violin()

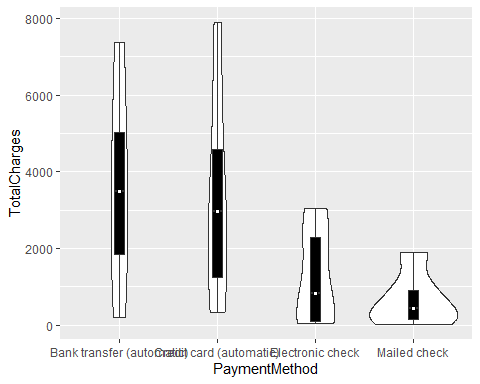


telecom+geom\_violin(adjust=2)

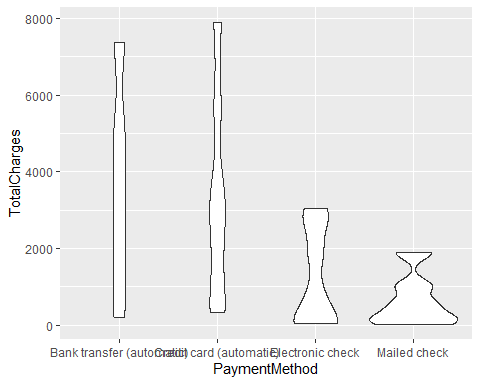


telecom+geom\_violin()+geom\_boxplot(width=.1,fill="black",outlier.colour=NA)+stat\_summary(fun.y=median,geom="point",fill="white",shape=21,size=1.5)

## Warning: The `fun.y` argument of `stat\_summary()` is deprecated as of ggplot2 3.3.0.  
## ℹ Please use the `fun` argument instead.  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last\_lifecycle\_warnings()` to see where this warning was  
## generated.



telecom+geom\_violin(adjust=.5)



str(telecom)

## List of 9  
## $ data : tibble [25 × 21] (S3: tbl\_df/tbl/data.frame)  
## ..$ customerID : chr [1:25] "7590-VHVEG" "5575-GNVDE" "3668-QPYBK" "7795-CFOCW" ...  
## ..$ gender : chr [1:25] "Female" "Male" "Male" "Male" ...  
## ..$ SeniorCitizen : num [1:25] 0 0 0 0 0 0 0 0 0 0 ...  
## ..$ Partner : chr [1:25] "Yes" "No" "No" "No" ...  
## ..$ Dependents : chr [1:25] "No" "No" "No" "No" ...  
## ..$ tenure : num [1:25] 1 34 2 45 2 8 22 10 28 62 ...  
## ..$ PhoneService : chr [1:25] "No" "Yes" "Yes" "No" ...  
## ..$ MultipleLines : chr [1:25] "No phone service" "No" "No" "No phone service" ...  
## ..$ InternetService : chr [1:25] "DSL" "DSL" "DSL" "DSL" ...  
## ..$ OnlineSecurity : chr [1:25] "No" "Yes" "Yes" "Yes" ...  
## ..$ OnlineBackup : chr [1:25] "Yes" "No" "Yes" "No" ...  
## ..$ DeviceProtection: chr [1:25] "No" "Yes" "No" "Yes" ...  
## ..$ TechSupport : chr [1:25] "No" "No" "No" "Yes" ...  
## ..$ StreamingTV : chr [1:25] "No" "No" "No" "No" ...  
## ..$ StreamingMovies : chr [1:25] "No" "No" "No" "No" ...  
## ..$ Contract : chr [1:25] "Month-to-month" "One year" "Month-to-month" "One year" ...  
## ..$ PaperlessBilling: chr [1:25] "Yes" "No" "Yes" "No" ...  
## ..$ PaymentMethod : chr [1:25] "Electronic check" "Mailed check" "Mailed check" "Bank transfer (automatic)" ...  
## ..$ MonthlyCharges : num [1:25] 29.9 57 53.9 42.3 70.7 ...  
## ..$ TotalCharges : num [1:25] 29.9 1889.5 108.2 1840.8 151.7 ...  
## ..$ Churn : chr [1:25] "No" "No" "Yes" "No" ...  
## $ layers : list()  
## $ scales :Classes 'ScalesList', 'ggproto', 'gg' <ggproto object: Class ScalesList, gg>  
## add: function  
## clone: function  
## find: function  
## get\_scales: function  
## has\_scale: function  
## input: function  
## n: function  
## non\_position\_scales: function  
## scales: NULL  
## super: <ggproto object: Class ScalesList, gg>   
## $ mapping :List of 2  
## ..$ x: language ~PaymentMethod  
## .. ..- attr(\*, ".Environment")=<environment: R\_GlobalEnv>   
## ..$ y: language ~TotalCharges  
## .. ..- attr(\*, ".Environment")=<environment: R\_GlobalEnv>   
## ..- attr(\*, "class")= chr "uneval"  
## $ theme : list()  
## $ coordinates:Classes 'CoordCartesian', 'Coord', 'ggproto', 'gg' <ggproto object: Class CoordCartesian, Coord, gg>  
## aspect: function  
## backtransform\_range: function  
## clip: on  
## default: TRUE  
## distance: function  
## expand: TRUE  
## is\_free: function  
## is\_linear: function  
## labels: function  
## limits: list  
## modify\_scales: function  
## range: function  
## render\_axis\_h: function  
## render\_axis\_v: function  
## render\_bg: function  
## render\_fg: function  
## setup\_data: function  
## setup\_layout: function  
## setup\_panel\_guides: function  
## setup\_panel\_params: function  
## setup\_params: function  
## train\_panel\_guides: function  
## transform: function  
## super: <ggproto object: Class CoordCartesian, Coord, gg>   
## $ facet :Classes 'FacetNull', 'Facet', 'ggproto', 'gg' <ggproto object: Class FacetNull, Facet, gg>  
## compute\_layout: function  
## draw\_back: function  
## draw\_front: function  
## draw\_labels: function  
## draw\_panels: function  
## finish\_data: function  
## init\_scales: function  
## map\_data: function  
## params: list  
## setup\_data: function  
## setup\_params: function  
## shrink: TRUE  
## train\_scales: function  
## vars: function  
## super: <ggproto object: Class FacetNull, Facet, gg>   
## $ plot\_env :<environment: R\_GlobalEnv>   
## $ labels :List of 2  
## ..$ x: chr "PaymentMethod"  
## ..$ y: chr "TotalCharges"  
## - attr(\*, "class")= chr [1:2] "gg" "ggplot"

seq(telecom)

## [1] 1 2 3 4 5 6 7 8 9

names(telecom)

## [1] "data" "layers" "scales" "mapping" "theme"   
## [6] "coordinates" "facet" "plot\_env" "labels"

class(telecom)

## [1] "gg" "ggplot"

typeof(telecom)

## [1] "list"

ls()

## [1] "myfun" "telecom"

unique(telecom$PaymentMethod)

## NULL

nrow(telecom)

## NULL

ncol(telecom)

## NULL

cumsum(telecom$gender)

## numeric(0)

cumprod(telecom$tenure)

## numeric(0)

cummin(telecom$OnlineSecurity)

## numeric(0)

cummax(telecom$TotalCharges)

## numeric(0)

length(telecom)

## [1] 9

rev(telecom$Partner)

## NULL

rep(telecom)

## $data  
## # A tibble: 25 × 21  
## customerID gender SeniorCitizen Partner Dependents tenure PhoneService  
## <chr> <chr> <dbl> <chr> <chr> <dbl> <chr>   
## 1 7590-VHVEG Female 0 Yes No 1 No   
## 2 5575-GNVDE Male 0 No No 34 Yes   
## 3 3668-QPYBK Male 0 No No 2 Yes   
## 4 7795-CFOCW Male 0 No No 45 No   
## 5 9237-HQITU Female 0 No No 2 Yes   
## 6 9305-CDSKC Female 0 No No 8 Yes   
## 7 1452-KIOVK Male 0 No Yes 22 Yes   
## 8 6713-OKOMC Female 0 No No 10 No   
## 9 7892-POOKP Female 0 Yes No 28 Yes   
## 10 6388-TABGU Male 0 No Yes 62 Yes   
## # ℹ 15 more rows  
## # ℹ 14 more variables: MultipleLines <chr>, InternetService <chr>,  
## # OnlineSecurity <chr>, OnlineBackup <chr>, DeviceProtection <chr>,  
## # TechSupport <chr>, StreamingTV <chr>, StreamingMovies <chr>,  
## # Contract <chr>, PaperlessBilling <chr>, PaymentMethod <chr>,  
## # MonthlyCharges <dbl>, TotalCharges <dbl>, Churn <chr>  
##   
## $layers  
## list()  
##   
## $scales  
## <ggproto object: Class ScalesList, gg>  
## add: function  
## clone: function  
## find: function  
## get\_scales: function  
## has\_scale: function  
## input: function  
## n: function  
## non\_position\_scales: function  
## scales: NULL  
## super: <ggproto object: Class ScalesList, gg>  
##   
## $mapping  
## Aesthetic mapping:   
## \* `x` -> `PaymentMethod`  
## \* `y` -> `TotalCharges`  
##   
## $theme  
## list()  
##   
## $coordinates  
## <ggproto object: Class CoordCartesian, Coord, gg>  
## aspect: function  
## backtransform\_range: function  
## clip: on  
## default: TRUE  
## distance: function  
## expand: TRUE  
## is\_free: function  
## is\_linear: function  
## labels: function  
## limits: list  
## modify\_scales: function  
## range: function  
## render\_axis\_h: function  
## render\_axis\_v: function  
## render\_bg: function  
## render\_fg: function  
## setup\_data: function  
## setup\_layout: function  
## setup\_panel\_guides: function  
## setup\_panel\_params: function  
## setup\_params: function  
## train\_panel\_guides: function  
## transform: function  
## super: <ggproto object: Class CoordCartesian, Coord, gg>  
##   
## $facet  
## <ggproto object: Class FacetNull, Facet, gg>  
## compute\_layout: function  
## draw\_back: function  
## draw\_front: function  
## draw\_labels: function  
## draw\_panels: function  
## finish\_data: function  
## init\_scales: function  
## map\_data: function  
## params: list  
## setup\_data: function  
## setup\_params: function  
## shrink: TRUE  
## train\_scales: function  
## vars: function  
## super: <ggproto object: Class FacetNull, Facet, gg>  
##   
## $plot\_env  
## <environment: R\_GlobalEnv>  
##   
## $labels  
## $labels$x  
## [1] "PaymentMethod"  
##   
## $labels$y  
## [1] "TotalCharges"

sd(telecom$MonthlyCharges)

## [1] NA

median(telecom$MonthlyCharges)

## NULL

mean(telecom$MonthlyCharges)

## Warning in mean.default(telecom$MonthlyCharges): argument is not numeric or  
## logical: returning NA

## [1] NA

sum(telecom$TotalCharges)

## [1] 0

telecom$customerID

## NULL

formatC(telecom$Dependents)

## character(0)

head(telecom)

## $data  
## # A tibble: 25 × 21  
## customerID gender SeniorCitizen Partner Dependents tenure PhoneService  
## <chr> <chr> <dbl> <chr> <chr> <dbl> <chr>   
## 1 7590-VHVEG Female 0 Yes No 1 No   
## 2 5575-GNVDE Male 0 No No 34 Yes   
## 3 3668-QPYBK Male 0 No No 2 Yes   
## 4 7795-CFOCW Male 0 No No 45 No   
## 5 9237-HQITU Female 0 No No 2 Yes   
## 6 9305-CDSKC Female 0 No No 8 Yes   
## 7 1452-KIOVK Male 0 No Yes 22 Yes   
## 8 6713-OKOMC Female 0 No No 10 No   
## 9 7892-POOKP Female 0 Yes No 28 Yes   
## 10 6388-TABGU Male 0 No Yes 62 Yes   
## # ℹ 15 more rows  
## # ℹ 14 more variables: MultipleLines <chr>, InternetService <chr>,  
## # OnlineSecurity <chr>, OnlineBackup <chr>, DeviceProtection <chr>,  
## # TechSupport <chr>, StreamingTV <chr>, StreamingMovies <chr>,  
## # Contract <chr>, PaperlessBilling <chr>, PaymentMethod <chr>,  
## # MonthlyCharges <dbl>, TotalCharges <dbl>, Churn <chr>  
##   
## $layers  
## list()  
##   
## $scales  
## <ggproto object: Class ScalesList, gg>  
## add: function  
## clone: function  
## find: function  
## get\_scales: function  
## has\_scale: function  
## input: function  
## n: function  
## non\_position\_scales: function  
## scales: NULL  
## super: <ggproto object: Class ScalesList, gg>  
##   
## $mapping  
## Aesthetic mapping:   
## \* `x` -> `PaymentMethod`  
## \* `y` -> `TotalCharges`  
##   
## $theme  
## list()  
##   
## $coordinates  
## <ggproto object: Class CoordCartesian, Coord, gg>  
## aspect: function  
## backtransform\_range: function  
## clip: on  
## default: TRUE  
## distance: function  
## expand: TRUE  
## is\_free: function  
## is\_linear: function  
## labels: function  
## limits: list  
## modify\_scales: function  
## range: function  
## render\_axis\_h: function  
## render\_axis\_v: function  
## render\_bg: function  
## render\_fg: function  
## setup\_data: function  
## setup\_layout: function  
## setup\_panel\_guides: function  
## setup\_panel\_params: function  
## setup\_params: function  
## train\_panel\_guides: function  
## transform: function  
## super: <ggproto object: Class CoordCartesian, Coord, gg>

tail(telecom)

## $mapping  
## Aesthetic mapping:   
## \* `x` -> `PaymentMethod`  
## \* `y` -> `TotalCharges`  
##   
## $theme  
## list()  
##   
## $coordinates  
## <ggproto object: Class CoordCartesian, Coord, gg>  
## aspect: function  
## backtransform\_range: function  
## clip: on  
## default: TRUE  
## distance: function  
## expand: TRUE  
## is\_free: function  
## is\_linear: function  
## labels: function  
## limits: list  
## modify\_scales: function  
## range: function  
## render\_axis\_h: function  
## render\_axis\_v: function  
## render\_bg: function  
## render\_fg: function  
## setup\_data: function  
## setup\_layout: function  
## setup\_panel\_guides: function  
## setup\_panel\_params: function  
## setup\_params: function  
## train\_panel\_guides: function  
## transform: function  
## super: <ggproto object: Class CoordCartesian, Coord, gg>  
##   
## $facet  
## <ggproto object: Class FacetNull, Facet, gg>  
## compute\_layout: function  
## draw\_back: function  
## draw\_front: function  
## draw\_labels: function  
## draw\_panels: function  
## finish\_data: function  
## init\_scales: function  
## map\_data: function  
## params: list  
## setup\_data: function  
## setup\_params: function  
## shrink: TRUE  
## train\_scales: function  
## vars: function  
## super: <ggproto object: Class FacetNull, Facet, gg>  
##   
## $plot\_env  
## <environment: R\_GlobalEnv>  
##   
## $labels  
## $labels$x  
## [1] "PaymentMethod"  
##   
## $labels$y  
## [1] "TotalCharges"

telecom$InternetService

## NULL

is.data.frame(telecom)

## [1] FALSE

is.name(telecom)

## [1] FALSE

identity(telecom$gender)

## NULL

binom.test(29,200,.21)

##   
## Exact binomial test  
##   
## data: 29 and 200  
## number of successes = 29, number of trials = 200, p-value = 0.02374  
## alternative hypothesis: true probability of success is not equal to 0.21  
## 95 percent confidence interval:  
## 0.09930862 0.20156150  
## sample estimates:  
## probability of success   
## 0.145

pbinom(2,5,.5)

## [1] 0.5

dbinom(0,5,.5)

## [1] 0.03125

prop.test(29,200,.21)

##   
## 1-sample proportions test with continuity correction  
##   
## data: 29 out of 200, null probability 0.21  
## X-squared = 4.7092, df = 1, p-value = 0.03  
## alternative hypothesis: true p is not equal to 0.21  
## 95 percent confidence interval:  
## 0.1007793 0.2032735  
## sample estimates:  
## p   
## 0.145

diff(telecom$TotalCharges)

## NULL

pmax(telecom$TotalCharges)

## NULL

pmin(telecom$TotalCharges)

## NULL

summary(telecom$customerID)

## Length Class Mode   
## 0 NULL NULL

summary(telecom$gender)

## Length Class Mode   
## 0 NULL NULL

summary(telecom$SeniorCitizen)

## Length Class Mode   
## 0 NULL NULL

summary(telecom$Partner)

## Length Class Mode   
## 0 NULL NULL

summary(telecom$Dependents)

## Length Class Mode   
## 0 NULL NULL

summary(telecom$MultipleLines)

## Length Class Mode   
## 0 NULL NULL

dim(telecom)

## NULL

summary(telecom$Contract)

## Length Class Mode   
## 0 NULL NULL

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

summary(cars)

## speed dist   
## Min. : 4.0 Min. : 2.00   
## 1st Qu.:12.0 1st Qu.: 26.00   
## Median :15.0 Median : 36.00   
## Mean :15.4 Mean : 42.98   
## 3rd Qu.:19.0 3rd Qu.: 56.00   
## Max. :25.0 Max. :120.00

## Including Plots

You can also embed plots, for example:



Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.