

# Visualización de datos en R

*joseangeldiazg*

*November 22, 2017*

## Ejercicio 1:

**Plot distance against stretch.**

Vamos a crear el dataframe con los datos de prueba:

```
stretch<-c(46,54,48,50,44,42,52)
distance<-c(148,182,173,166,109,141,166)
data<-data.frame(strech,distance)
```

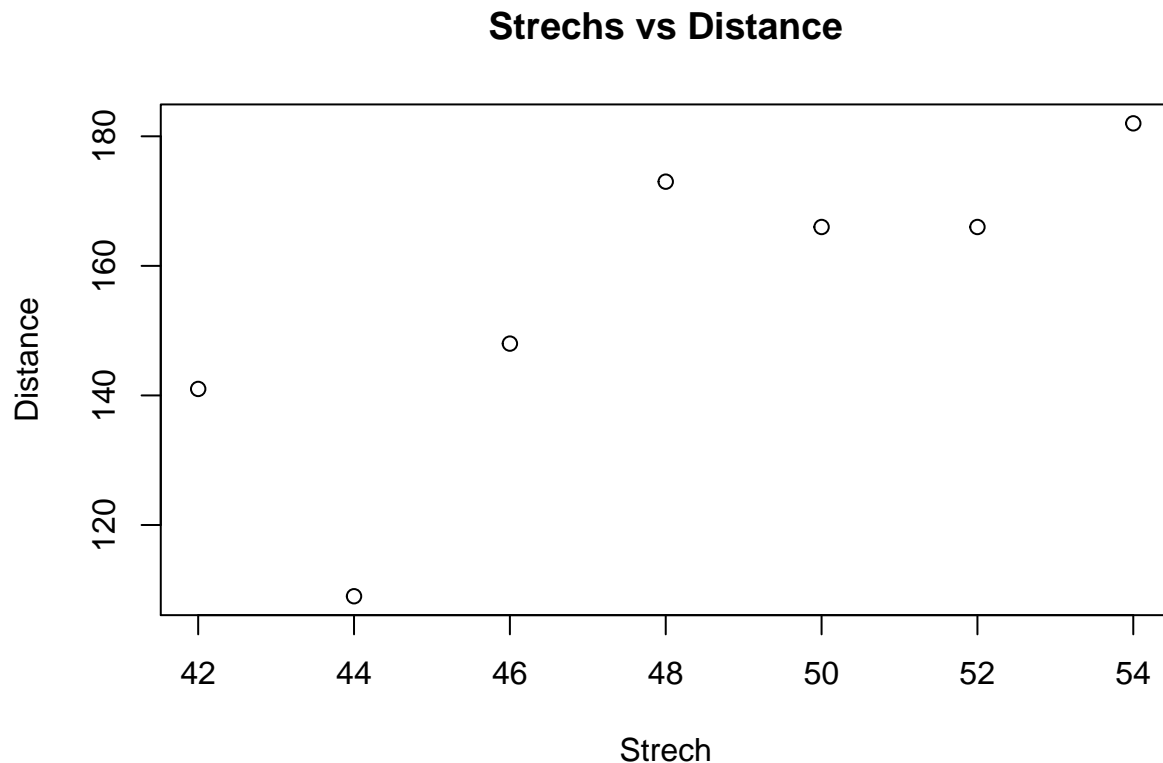
Visualizamos los datos creados:

```
data

##   stretch distance
## 1      46      148
## 2      54      182
## 3      48      173
## 4      50      166
## 5      44      109
## 6      42      141
## 7      52      166
```

Creamos el gráfico

```
plot(data$stretch, data$distance, main="Strechs vs Distance", xlab="Strech", ylab="Distance")
```



## Ejercicio 2:

The table on the right have ten observations, taken during the years 1970-79, are on October snow cover for Eurasia (snow cover is in millions of square kilometers):

- Plot snow.cover versus year:
- Plot a histogram of the snow.cover values:

Vamos a crear los datos, para ello procedemos como antes creando dos vectores y uniéndolos en un dataframe.

```
year<-c(1970,1971,1972,1973,1974,1975,1976,1977,1978,1979)
snow.cover<-c(6.5,12.0,14.9,10.0,10.7,7.9,21.9,12.5,14.5,9.2)
year.snow<-data.frame(year, snow.cover)
```

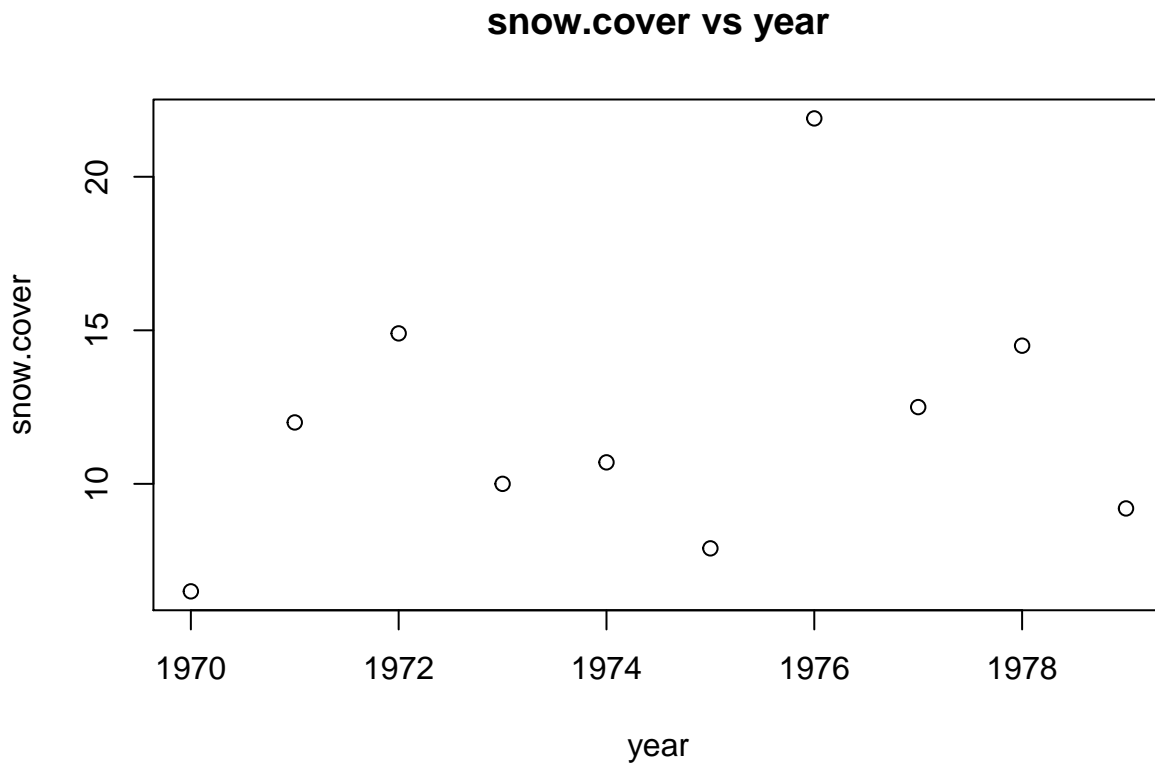
Mostramos el dataframe creado para ver si tenemos los daos como deseamos.

```
year.snow
```

```
##   year snow.cover
## 1  1970         6.5
## 2  1971        12.0
## 3  1972        14.9
## 4  1973        10.0
## 5  1974        10.7
## 6  1975         7.9
## 7  1976        21.9
## 8  1977        12.5
## 9  1978        14.5
## 10 1979         9.2
```

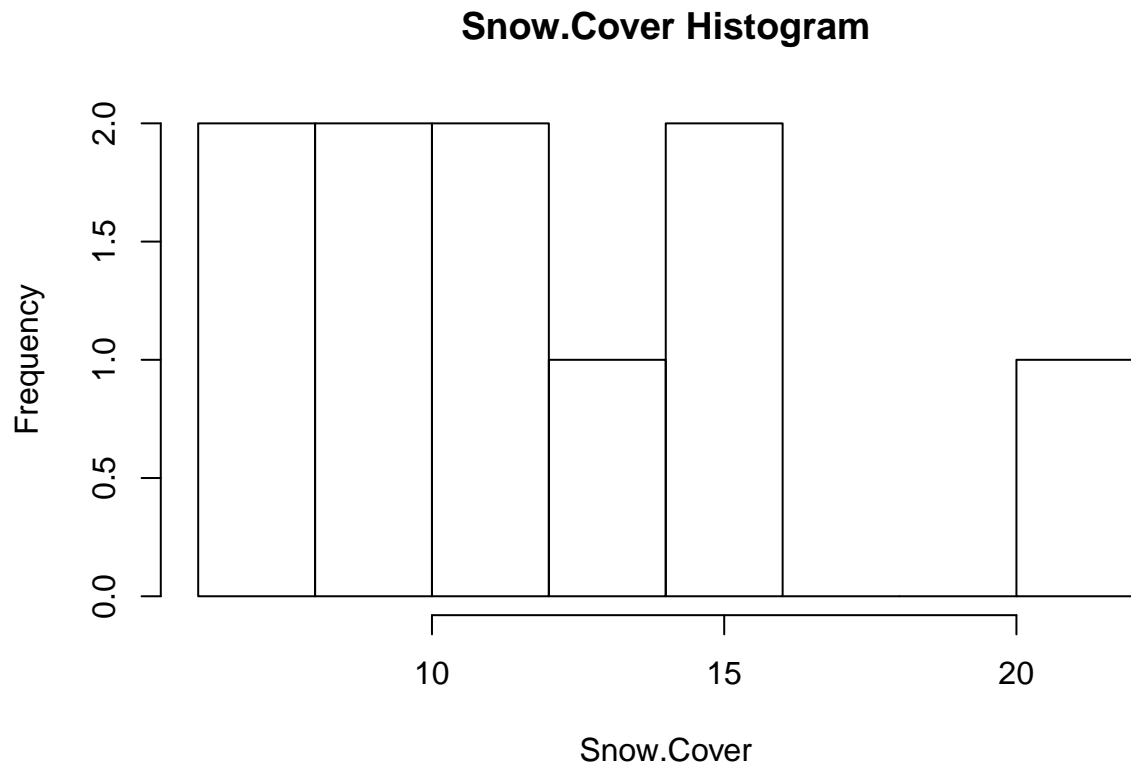
Pintamos el gráfico de **snow.cover** vs **year**, con lo que tendríamos hecho el primer apartado:

```
plot( year.snow$year,year.snow$snow.cover, main="snow.cover vs year", xlab="year", ylab="snow.cover" )
```



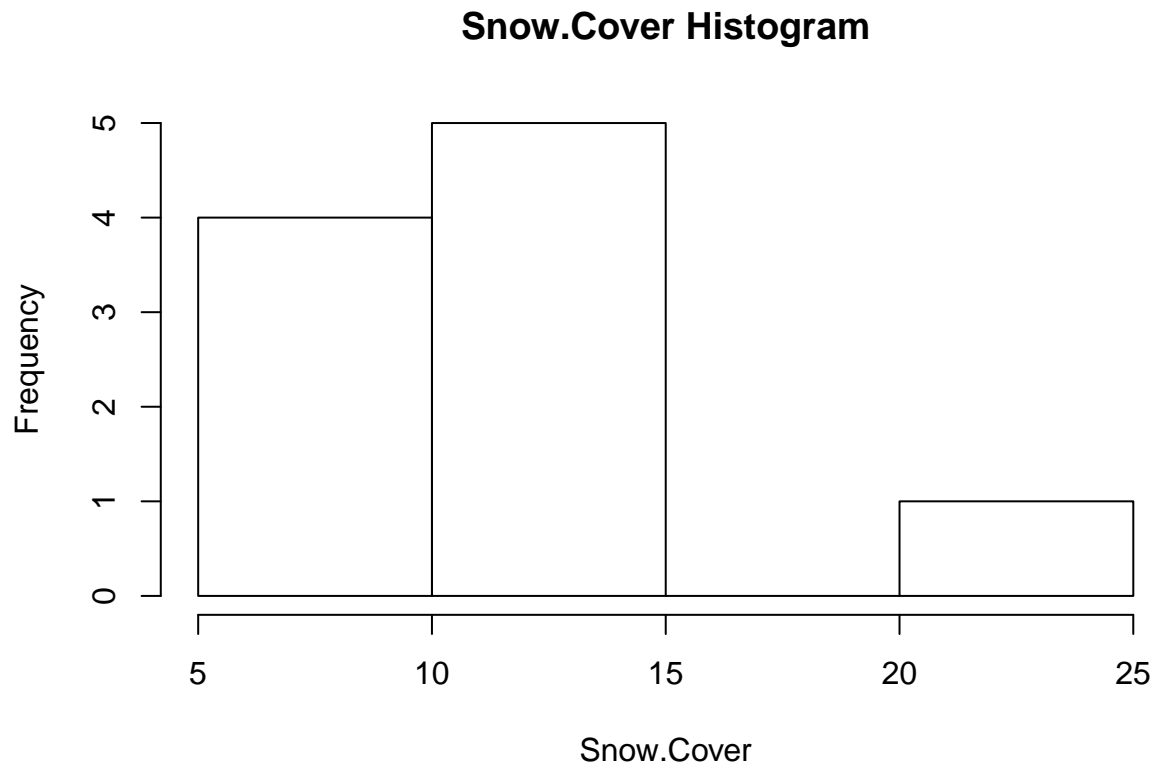
El siguiente punto pasa por la elaboración de un histograma, para ello podemos usar **hist()**

```
hist(snow.cover, main="Snow.Cover Histogram", xlab="Snow.Cover", breaks=10)
```



Vamos a probar con menos cortes:

```
hist(snow.cover, main="Snow.Cover Histogram", xlab="Snow.Cover", breaks=5)
```



A mi juicio este nos ofrece mejores resultados ya que podemos ver que no hay ninguno entre 15 y 20, que el mínimo está por encima de 5, y el máximo es  $\leq$  que 25, además como que la mayor parte de los años el

Snow.Cover ha estado situado entre 10 y 15.

## Ejercicio 3

Given the data in NY.xls

- (Optional)conver to F to oC and in to mm.
- PlotYear vs Warmest Minimum Temperature
- PlotYear vs Warmest Minimum Temperature and Coldest Minimum Temperature. Don't forget to add a legend!

El primer paso es cargar los datos para ello:

```
library(readxl)
NY <- read_excel("datasets/NY.xls")
View(NY)
```

El primer paso es convertir F a Grados e in a mm, para ello primero creamos dos funciones que realizan lo deseado:

```
ftoc<-function(gradosf)
{
  return ((gradosf-32)*(5/9))
}

intomm<-function(invalue)
{
  return (as.numeric(invalue)*25.4)
}
```

Ahora usamos la función para cada columna que queremos cambiar:

```
NY[2:8]<-apply(NY[2:8],2,ftoc)
```

Y por último cambiamos los elementos en in por mm:

```
NY[9:12]<-apply(NY[9:12],2,intomm)
```

```
## Warning in FUN(newX[, i], ...): NAs introducidos por coerción
```

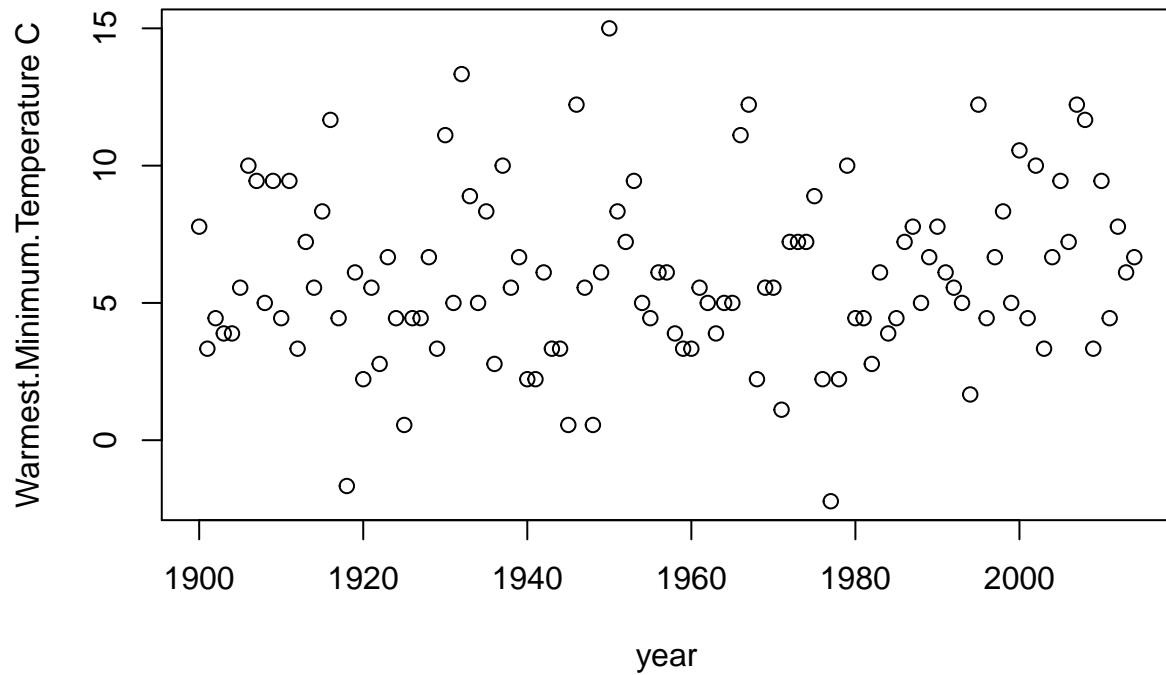
```
## Warning in FUN(newX[, i], ...): NAs introducidos por coerción
```

```
ny.dataframe<-data.frame(NY)
```

Ahora pintamos el gráfico como hemos hecho en puntos anteriores:

```
plot(ny.dataframe$Year, ny.dataframe$Warmest.Minimum.Temperature..F. , main="Warmest.Minimum.Temperature..F.")
```

## Warmest.Minimum.Temperature vs year

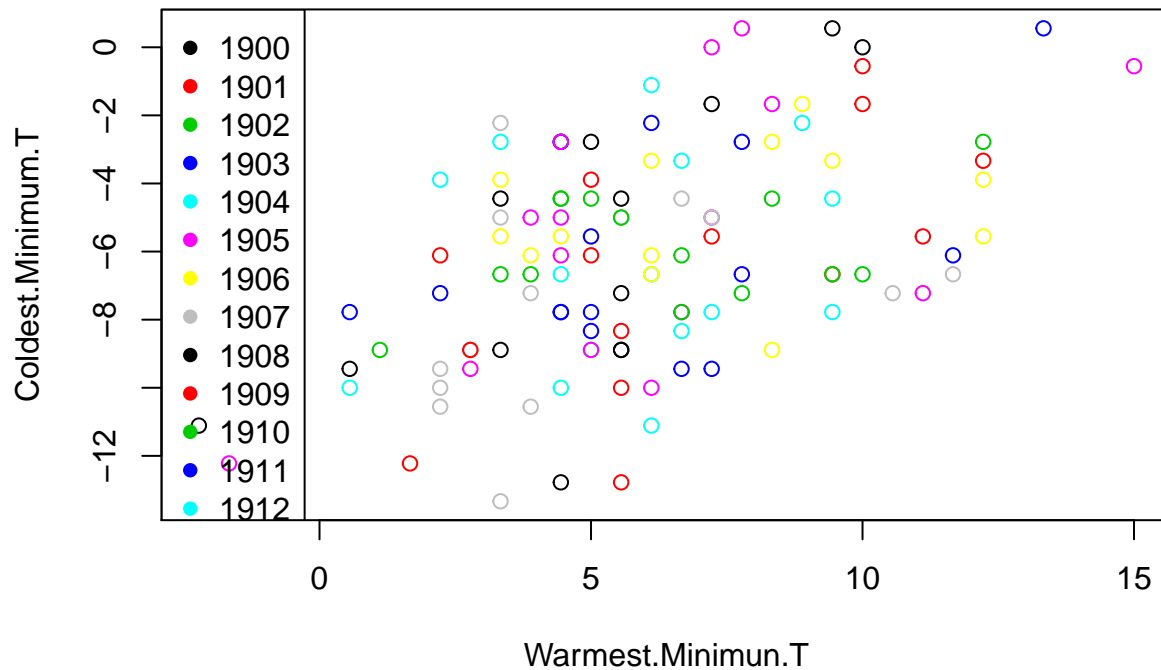


Por último pintamos el gráfico del punto 3 con Plot Year vs Warmest Minimum Temperature and Coldest Minimum Temperature

```
ny.dataframe$FactorYear<-as.factor(ny.dataframe$Year)
```

```
plot(ny.dataframe$Warmest.Minimum.Temperature..F., ny.dataframe$Coldest.Minimum.Temperature..F. , col=n  
legend("topleft", legend=levels(ny.dataframe$FactorYear), col=1:length(levels(ny.dataframe$FactorYear))
```

## Warmest.Minimum.Temperature vs Coldest.Minimum.Temperature



Como podemos ver, el gráfico no aporta mucha información, por lo que vamos a reducir la dimensionalidad de nuestra variable para el año, creando una variable categórica por ejemplo para los cuartos de siglo, lo que por ejemplo sería útil para estudiar tendencias por ejemplo en la temperatura.

```
obtieneSiglo<-function(year)
{
  year<-as.numeric(year)
  print(year)
  if(year >= 1900 && year<1925)
  {
    return("1º Cuarto Siglo XX")
  }
  else if (year >= 1925 && year<1950)
  {
    return("2º Cuarto Siglo XX")
  }
  else if (year >= 1950 && year<1975)
  {
    return("3º Cuarto Siglo XX")
  }
  else if (year >= 1975 && year<2000)
  {
    return("Final Siglo XX")
  }
  else
  {
    return("Siglo XXI")
  }
}
```

```
ny.dataframe$CategoricYear<-lapply(ny.dataframe$Year,obtieneSiglo)
```

```
## [1] 2014
## [1] 2013
## [1] 2012
## [1] 2011
## [1] 2010
## [1] 2009
## [1] 2008
## [1] 2007
## [1] 2006
## [1] 2005
## [1] 2004
## [1] 2003
## [1] 2002
## [1] 2001
## [1] 2000
## [1] 1999
## [1] 1998
## [1] 1997
## [1] 1996
## [1] 1995
## [1] 1994
## [1] 1993
## [1] 1992
## [1] 1991
## [1] 1990
## [1] 1989
## [1] 1988
## [1] 1987
## [1] 1986
## [1] 1985
## [1] 1984
## [1] 1983
## [1] 1982
## [1] 1981
## [1] 1980
## [1] 1979
## [1] 1978
## [1] 1977
## [1] 1976
## [1] 1975
## [1] 1974
## [1] 1973
## [1] 1972
## [1] 1971
## [1] 1970
## [1] 1969
## [1] 1968
## [1] 1967
## [1] 1966
## [1] 1965
## [1] 1964
## [1] 1963
```



## [1] 1962  
## [1] 1961  
## [1] 1960  
## [1] 1959  
## [1] 1958  
## [1] 1957  
## [1] 1956  
## [1] 1955  
## [1] 1954  
## [1] 1953  
## [1] 1952  
## [1] 1951  
## [1] 1950  
## [1] 1949  
## [1] 1948  
## [1] 1947  
## [1] 1946  
## [1] 1945  
## [1] 1944  
## [1] 1943  
## [1] 1942  
## [1] 1941  
## [1] 1940  
## [1] 1939  
## [1] 1938  
## [1] 1937  
## [1] 1936  
## [1] 1935  
## [1] 1934  
## [1] 1933  
## [1] 1932  
## [1] 1931  
## [1] 1930  
## [1] 1929  
## [1] 1928  
## [1] 1927  
## [1] 1926  
## [1] 1925  
## [1] 1924  
## [1] 1923  
## [1] 1922  
## [1] 1921  
## [1] 1920  
## [1] 1919  
## [1] 1918  
## [1] 1917  
## [1] 1916  
## [1] 1915  
## [1] 1914  
## [1] 1913  
## [1] 1912  
## [1] 1911  
## [1] 1910  
## [1] 1909

```
## [1] 1908
## [1] 1907
## [1] 1906
## [1] 1905
## [1] 1904
## [1] 1903
## [1] 1902
## [1] 1901
## [1] 1900
```

Ahora solo tenemos que dibujar nuestro gráfico con la variable categórica creada, pero antes debemos pasarla a factor.

```
ny.dataframe$CategoricYear<-unlist(ny.dataframe$CategoricYear)
ny.dataframe$FactorCategoricYear<-as.factor(ny.dataframe$CategoricYear)

plot(ny.dataframe$Warmest.Minimum.Temperature..F., ny.dataframe$Coldest.Minimum.Temperature..F. , col=ny
legend("topleft", legend=levels(ny.dataframe$FactorCategoricYear),col=1:length(levels(ny.dataframe$FactorCategoricYear))
```

## Warmest.Minimum.Temperature vs Coldest.Minimum.Temperature

