

```
## Warning in readChar(con, 5L, useBytes = TRUE): cannot open compressed
file '.RData', probable reason 'No such file or directory'
## Error in readChar(con, 5L, useBytes = TRUE): cannot open the connection
```

The first step in analyzing water chemistry data for CWA compliance is reading it into the analytical software.

```
carlin <- read.csv("./carlin.csv", header = TRUE, sep = ",", stringsAsFactors = F)
```

Next, check that the data are what you expect to see and convert dates from factors.

```
carlin$sampdate <- as.Date(carlin$sampdate)
str(carlin)

## 'data.frame': 209 obs. of 52 variables:
## $ siteid : int 10321000 10321000 10321000 10321000 10321000 10321000 10321000 10321000 10321000 10321000
## $ sampdate: Date, format: "1965-10-01" "1965-11-01" ...
## $ Temp.h2o: num NA NA NA NA NA NA NA NA NA NA NA ...
## $ Temp.air: num NA NA NA NA NA NA NA NA NA NA NA ...
## $ Disc.cfs: num 98 141 128 122 97 108 384 701 612 541 ...
## $ Turb : int NA NA NA NA NA NA NA NA NA NA NA ...
## $ SC : int 490 507 564 506 551 506 516 628 470 423 ...
## $ DO : num NA NA NA NA NA NA NA NA NA NA NA ...
## $ pH : num 8.4 8.3 8.1 8.2 8.5 8.2 7.7 8.1 8 8 ...
## $ ANC : int NA 213 238 212 241 216 201 235 189 172 ...
## $ HC03 : int 238 252 290 259 276 263 245 286 231 210 ...
## $ C03 : int 4 4 0 0 9 0 0 0 0 0 ...
## $ Alk : int NA NA NA NA NA NA NA NA NA NA NA ...
## $ Hard : num 150 170 190 170 200 180 160 190 160 140 ...
## $ TDS : int NA NA 359 321 NA NA NA NA 301 NA ...
## $ TSS : int NA NA NA NA NA NA NA NA NA NA NA ...
## $ N.tot : num NA NA NA NA NA NA NA NA NA NA NA ...
## $ N.org : num NA NA NA NA NA NA NA NA NA NA NA ...
## $ NH4 : num NA NA NA NA NA NA NA NA NA NA NA ...
## $ NO3 : num NA NA NA NA NA NA NA NA NA NA NA ...
## $ NO2 : num NA NA NA NA NA NA NA NA NA NA NA ...
## $ PO4 : num NA NA NA NA NA NA NA NA NA NA NA ...
## $ C : num NA NA NA NA NA NA NA NA NA NA NA ...
## $ Ca : num 40 50 56 50 58 52 46 51 43 40 ...
## $ Mg : num 12 11 13 11 13 11 11 16 12 10 ...
## $ Na : num 46 42 47 45 48 42 46 61 40 34 ...
## $ K : num NA NA 6.4 6.1 NA NA 8 8.5 6.3 5.4 ...
## $ Cl : num NA NA 16 16 NA NA NA NA 14 NA ...
## $ SO4 : num NA NA 37 34 NA NA NA NA 37 NA ...
```

```
## $ F      : num  NA NA 0.5 0.6 NA NA NA NA 0.5 NA ...
## $ Si     : num  NA NA 40 31 NA NA NA NA 34 NA ...
## $ As     : int   NA NA NA NA NA NA NA NA NA NA ...
## $ Ba     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Be     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Cd     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Cr     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Co     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Cu     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Fe     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Pb     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Mn     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Mo     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Ni     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Ag     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Sr     : int   NA NA NA NA NA NA NA NA NA NA ...
## $ V      : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Zi     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Al     : int   NA NA NA NA NA NA NA NA NA NA ...
## $ Li     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Se     : int   NA NA NA NA NA NA NA NA NA NA ...
## $ CN     : num  NA NA NA NA NA NA NA NA NA NA ...
## $ Hg     : num  NA NA NA NA NA NA NA NA NA NA ...
```

The above table shows the structure of the data in a R data.frame format.  
Another perspective on the data is the summary of each of the data columns.

```
summary(carlin)

##      siteid      sampdate      Temp.h2o      Temp.air
## Min.   :10321000  Min.   :1965-10-01  Min.   : 0.00  Min.   : -7.00
## 1st Qu.:10321000  1st Qu.:1968-04-01  1st Qu.: 4.00  1st Qu.:  6.00
## Median :10321000  Median :1981-01-28  Median :13.00  Median :16.75
## Mean   :10321000  Mean   :1980-11-09  Mean   :12.05  Mean   :14.83
## 3rd Qu.:10321000  3rd Qu.:1989-05-31  3rd Qu.:19.50  3rd Qu.:22.25
## Max.   :10321000  Max.   :2015-06-26  Max.   :26.00  Max.   :34.50
##                                     NA's   :68      NA's   :131
##      Disc.cfs      Turb      SC      DO
## Min.   :  3.9  Min.   :  1.00  Min.   :300.0  Min.   :  3.00
## 1st Qu.: 37.0  1st Qu.:  5.00  1st Qu.:416.5  1st Qu.:  8.55
## Median :120.5  Median :12.50  Median :479.0  Median :  9.80
## Mean   :363.8  Mean   :29.49  Mean   :477.7  Mean   :10.16
## 3rd Qu.:365.5  3rd Qu.:31.00  3rd Qu.:532.8  3rd Qu.:12.10
## Max.   :8130.0  Max.   :640.00  Max.   :675.0  Max.   :15.20
## NA's   :11     NA's   :97     NA's   :3     NA's   :98
##      pH      ANC      HC03      C03
```

##	Min.	:7.600	Min.	:109.0	Min.	:155.0	Min.	: 0.000
##	1st Qu.:	8.200	1st Qu.:	174.0	1st Qu.:	227.0	1st Qu.:	0.000
##	Median	:8.400	Median	:200.0	Median	:240.0	Median	: 3.000
##	Mean	:8.362	Mean	:196.6	Mean	:243.4	Mean	: 4.565
##	3rd Qu.:	8.600	3rd Qu.:	216.2	3rd Qu.:	260.0	3rd Qu.:	7.000
##	Max.	:8.900	Max.	:280.0	Max.	:340.0	Max.	:42.000
##	NA's	:29	NA's	:41	NA's	:124	NA's	:124
##	Alk		Hard		TDS		TSS	
##	Min.	:131.0	Min.	: 79.9	Min.	:175.0	Min.	:43.00
##	1st Qu.:	163.5	1st Qu.:	148.0	1st Qu.:	259.0	1st Qu.:	70.50
##	Median	:181.5	Median	:162.0	Median	:295.0	Median	:83.00
##	Mean	:184.9	Mean	:162.7	Mean	:297.9	Mean	:79.66
##	3rd Qu.:	198.5	3rd Qu.:	180.0	3rd Qu.:	334.5	3rd Qu.:	89.00
##	Max.	:270.0	Max.	:219.0	Max.	:450.0	Max.	:99.00
##	NA's	:163	NA's	:34	NA's	:86	NA's	:147
##	N.tot		N.org		NH4		N03	
##	Min.	:0.250	Min.	:0.1600	Min.	:0.00000	Min.	:0.02000
##	1st Qu.:	0.500	1st Qu.:	0.3800	1st Qu.:	0.01000	1st Qu.:	0.05000
##	Median	:0.620	Median	:0.4900	Median	:0.02000	Median	:0.07850
##	Mean	:0.769	Mean	:0.6102	Mean	:0.04373	Mean	:0.08338
##	3rd Qu.:	1.000	3rd Qu.:	0.7700	3rd Qu.:	0.05750	3rd Qu.:	0.10000
##	Max.	:2.600	Max.	:2.4000	Max.	:0.33000	Max.	:0.62000
##	NA's	:111	NA's	:112	NA's	:99	NA's	:141
##	N02		P04		C		Ca	
##	Min.	:0.0100	Min.	:0.0310	Min.	: 4.000	Min.	:23.00
##	1st Qu.:	0.0100	1st Qu.:	0.0310	1st Qu.:	5.050	1st Qu.:	42.00
##	Median	:0.0100	Median	:0.0920	Median	: 6.950	Median	:46.00
##	Mean	:0.0106	Mean	:0.1489	Mean	: 7.642	Mean	:46.25
##	3rd Qu.:	0.0100	3rd Qu.:	0.1840	3rd Qu.:	9.325	3rd Qu.:	50.00
##	Max.	:0.0200	Max.	:1.1700	Max.	:13.000	Max.	:64.00
##	NA's	:142	NA's	:121	NA's	:197	NA's	:34
##	Mg		Na		K		Cl	
##	Min.	: 5.40	Min.	:19.00	Min.	: 3.700	Min.	: 6.90
##	1st Qu.:	10.00	1st Qu.:	33.00	1st Qu.:	5.800	1st Qu.:	14.00
##	Median	:12.00	Median	:40.00	Median	: 6.900	Median	:16.00
##	Mean	:11.41	Mean	:39.92	Mean	: 7.513	Mean	:16.68
##	3rd Qu.:	13.00	3rd Qu.:	47.00	3rd Qu.:	7.900	3rd Qu.:	20.00
##	Max.	:17.00	Max.	:66.00	Max.	:96.000	Max.	:40.00
##	NA's	:34	NA's	:34	NA's	:46	NA's	:82
##	S04		F		Si		As	
##	Min.	:11.00	Min.	:0.1000	Min.	:15.00	Min.	: 3.000
##	1st Qu.:	26.00	1st Qu.:	0.4000	1st Qu.:	23.00	1st Qu.:	6.000
##	Median	:33.00	Median	:0.5000	Median	:27.00	Median	: 7.000
##	Mean	:33.89	Mean	:0.4758	Mean	:26.17	Mean	: 7.079
##	3rd Qu.:	41.00	3rd Qu.:	0.5250	3rd Qu.:	29.00	3rd Qu.:	8.000

##	Max.	:63.00	Max.	:1.3000	Max.	:40.00	Max.	:14.000
##	NA's	:82	NA's	:81	NA's	:81	NA's	:146
##		Ba		Be		Cd		Cr
##	Min.	: 47.40	Min.	:0.5000	Min.	:1.000	Min.	: 0.00
##	1st Qu.:	72.00	1st Qu.:	0.5000	1st Qu.:	1.000	1st Qu.:	1.00
##	Median :	90.00	Median :	0.5000	Median :	1.000	Median :	1.00
##	Mean :	88.87	Mean :	0.5792	Mean :	1.492	Mean :	1.61
##	3rd Qu.:	100.00	3rd Qu.:	0.5000	3rd Qu.:	1.000	3rd Qu.:	1.00
##	Max.	:140.00	Max.	:1.0000	Max.	:8.000	Max.	:10.00
##	NA's	:132	NA's	:161	NA's	:150	NA's	:150
##		Co		Cu		Fe		Pb
##	Min.	: 1.000	Min.	: 0.00	Min.	: 3.00	Min.	: 0.00
##	1st Qu.:	3.000	1st Qu.:	1.00	1st Qu.:	8.25	1st Qu.:	1.00
##	Median :	3.000	Median :	1.95	Median :	13.00	Median :	1.00
##	Mean :	3.408	Mean :	3.05	Mean :	22.39	Mean :	2.17
##	3rd Qu.:	3.000	3rd Qu.:	4.00	3rd Qu.:	27.00	3rd Qu.:	4.50
##	Max.	:12.000	Max.	:12.00	Max.	:130.00	Max.	:10.00
##	NA's	:133	NA's	:167	NA's	:139	NA's	:162
##		Mn		Mo		Ni		Ag
##	Min.	: 1.00	Min.	: 3.00	Min.	: 0.000	Min.	:0.0000
##	1st Qu.:	7.50	1st Qu.:	10.00	1st Qu.:	1.000	1st Qu.:	1.0000
##	Median :	10.00	Median :	10.00	Median :	1.000	Median :	1.0000
##	Mean :	13.95	Mean :	13.12	Mean :	1.516	Mean :	0.8816
##	3rd Qu.:	14.00	3rd Qu.:	10.00	3rd Qu.:	1.000	3rd Qu.:	1.0000
##	Max.	:160.00	Max.	:60.00	Max.	:10.000	Max.	:1.0000
##	NA's	:138	NA's	:146	NA's	:145	NA's	:133
##		Sr		V		Zi		Al
##	Min.	:150.0	Min.	: 6.000	Min.	: 3.00	Min.	: 5.00
##	1st Qu.:	260.0	1st Qu.:	6.000	1st Qu.:	3.00	1st Qu.:	10.00
##	Median :	365.0	Median :	6.000	Median :	7.00	Median :	10.00
##	Mean :	357.3	Mean :	6.298	Mean :	11.36	Mean :	25.92
##	3rd Qu.:	437.5	3rd Qu.:	6.000	3rd Qu.:	12.25	3rd Qu.:	24.00
##	Max.	:590.0	Max.	:10.000	Max.	:130.00	Max.	:180.00
##	NA's	:147	NA's	:147	NA's	:153	NA's	:145
##		Li		Se		CN		Hg
##	Min.	:13.00	Min.	:0.0000	Min.	:0.01	Min.	:0.0000
##	1st Qu.:	22.00	1st Qu.:	1.0000	1st Qu.:	0.01	1st Qu.:	0.1000
##	Median :	30.50	Median :	1.0000	Median :	0.01	Median :	0.1000
##	Mean :	32.93	Mean :	0.8961	Mean :	0.01	Mean :	0.1082
##	3rd Qu.:	41.00	3rd Qu.:	1.0000	3rd Qu.:	0.01	3rd Qu.:	0.1000
##	Max.	:67.00	Max.	:1.0000	Max.	:0.01	Max.	:0.5000
##	NA's	:147	NA's	:132	NA's	:181	NA's	:148

Plots of these distributions are the next step because the graphic conveys all the written information, plus much more insight into the data characteristics, more easily for decision-makers.

```
carlin.1 <- subset(carlin, select = siteid:C03)
carlin.2 <- subset(carlin, select = c(siteid, sampdate, Alk:Ca))
carlin.3 <- subset(carlin, select = c(siteid, sampdate, Mg:Cr))
carlin.4 <- subset(carlin, select = c(siteid, sampdate, Co:Hg))
```

To plot each chemical constituent's concentration as a function of collection date the data format needs to be reshaped from wide to long:

```
library(reshape)
carlin.1.melt <- melt(carlin.1, na.rm = F, id.vars = c("siteid", "sampdate"))
carlin.2.melt <- melt(carlin.2, na.rm = F, id.vars = c("siteid", "sampdate"))
carlin.3.melt <- melt(carlin.3, na.rm = F, id.vars = c("siteid", "sampdate"))
carlin.4.melt <- melt(carlin.4, na.rm = F, id.vars = c("siteid", "sampdate"))
```

The X-Y plots use only measured data; missing data are not included.

```
library(lattice)
xyplot(value ~ sampdate | variable, data = carlin.1.melt, rm.na = T)
```

