

STA5075: Practical 7

Jessica Stow (STWJES003@myuct.ac.za)

2025-02-04

Read the following data into R. Each time check that R has done the right thing. Leave the data as original as possible. 1. counts.xlsx (on Vula, Resources → data) 2. Tortoise data.xls, sheet 'Tortoise measurements' (on Vula) 3. Globular clusters: http://www.physics.mcmaster.ca/~harris/GCS_table.txt 4. Large-scale climatic index (MEI): <https://www.esrl.noaa.gov/psd/enso/mei/data/meiv2.data> 5. Voice Data from Singing the Vowel 'ooh': <http://www.statsci.org/data/general/ooh.txt> 6. Air quality data: air2.dat (on Vula) 7. 2D Pollen counts: 8.Red-cross-feb-aug-89-starts 20 Feb-1989.xls (on Vula) 8. wader counts.xls (on Vula)

```
# Set working directory
getwd()
```

```
## [1] "/Users/jessicastow/Documents/STA5075/sta5075"
```

```
# Clear workspace
rm(list = ls())
```

```
# Check workspace is clear
ls()
```

```
## character(0)
```

```
library(openxlsx) # for reading .xlsx files
library(readxl) # tidyverse - works both on .xls and .xlsx files
```

```
# can set stringsAsFactors = TRUE when reading in file
```

Question 1: counts.xlsx

```
df1 <- read.xlsx("counts.xlsx", sheet = 1)
```

```
# rename columns
```

```
colnames(df1)[1] <- "observation"
colnames(df1)[4] <- "comments"
```

```
# format as date
```

```
df1$day.month.year <- as.Date(df1$day.month.year, format = "%d %b %Y")
```

```
str(df1)
```

```
## 'data.frame':   15 obs. of  4 variables:
## $ observation   : num  1 2 3 4 5 6 7 8 9 10 ...
## $ day.month.year: Date, format: "2017-02-01" "2017-02-02" ...
## $ count         : num  6 7 3 4 5 NA 9 8 3 4 ...
## $ comments      : chr  NA NA NA NA ...
```

```
View(df1)
```

Question 2: Tortoise data.xls, sheet 'Tortoise measurements'

```
df2 <- read_excel("Tortoise data.xls", sheet = "Tortoise measurements")
```

```
## New names:  
## * `` -> `...12`
```

```
str(df2)
```

```
## tibble [243 x 12] (S3: tbl_df/tbl/data.frame)  
## $ No. : chr [1:243] "G1" "G2" "G3" "G4" ...  
## $ Sex : chr [1:243] "F" "M" "M" "M" ...  
## $ Habitat : chr [1:243] "Houses" "Open" "Open" "Open" ...  
## $ Behaviour : chr [1:243] "Passive" "Walking" "Walking" "Walking" ...  
## $ Weight : num [1:243] 1086 1311 1234 1119 1098 ...  
## $ Length : num [1:243] 178 238 213 205 200 160 221 179 179 190 ...  
## $ Plastron length: num [1:243] 134 163 152 155 148 128 164 146 138 145 ...  
## $ Gular length : num [1:243] 15 36 31 33 26 17 30 13 19 28 ...  
## $ Mid width : num [1:243] 114 126 125 117 115 ...  
## $ Curved width : num [1:243] 215 220 231 219 221 ...  
## $ Rear width : num [1:243] 99.5 125 132.2 129.5 113 ...  
## $ ...12 : chr [1:243] NA NA NA NA ...
```

```
# format as factors
```

```
df2$Sex <- as.factor(df2$Sex)  
df2$Habitat <- as.factor(df2$Habitat)  
df2$Behaviour <- as.factor(df2$Behaviour)
```

```
# rename last column
```

```
colnames(df2)[12] <- "Comments"
```

```
str(df2)
```

```
## tibble [243 x 12] (S3: tbl_df/tbl/data.frame)  
## $ No. : chr [1:243] "G1" "G2" "G3" "G4" ...  
## $ Sex : Factor w/ 3 levels "F","J","M": 1 3 3 3 3 1 3 1 1 3 ...  
## $ Habitat : Factor w/ 7 levels "Bush edge","Houses",...: 2 4 4 4 3 1 3 3 3 1 ...  
## $ Behaviour : Factor w/ 7 levels "Basking","Feeding",...: 6 7 7 7 2 6 2 6 6 6 ...  
## $ Weight : num [1:243] 1086 1311 1234 1119 1098 ...  
## $ Length : num [1:243] 178 238 213 205 200 160 221 179 179 190 ...  
## $ Plastron length: num [1:243] 134 163 152 155 148 128 164 146 138 145 ...  
## $ Gular length : num [1:243] 15 36 31 33 26 17 30 13 19 28 ...  
## $ Mid width : num [1:243] 114 126 125 117 115 ...  
## $ Curved width : num [1:243] 215 220 231 219 221 ...  
## $ Rear width : num [1:243] 99.5 125 132.2 129.5 113 ...  
## $ Comments : chr [1:243] NA NA NA NA ...
```

```
View(df2)
```

Question 3: Globular clusters http://www.physics.mcmaster.ca/~harris/GCS_table.txt

```
url_gcs <- "http://www.physics.mcmaster.ca/~harris/GCS_table.txt"

df3 <- read.table(url_gcs, sep = "", header = F, skip = 39, nrows = 422)
head(df3)
```

##	V1	V2	V3	V4	V5	V6	V7	V8	V9			
## 1	MilkyWay	nid	nd	nd	Sbc	nd	nd	geometry	0.000			
## 2	WLM	DDO-221	0.0328381	-15.4609285	Irr	1.00	0.10	TRGB/Cepheid	0.088			
## 3	NGC7814	nid	0.0541441	16.1455476	Sab	13.18	0.88	SBF	0.147			
## 4	NGC55	nid	0.2482245	-39.1965824	SBm	2.09	0.08	TRGB/Cepheid	0.036			
## 5	NGC147	nid	0.5533300	48.5086000	E5	0.76	0.10	TRGB/SBF	0.475			
## 6	NGC185	nid	0.6493737	48.3373957	E3	0.63	0.01	TRGB/SBF	0.500			
##	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21
## 1	-21.300	0.300	nd	nd	160.0	10	44	105.00	20.00	0.70	0.10	9.856
## 2	-14.840	0.200	nd	nd	1.0	0	50	nd	nd	nd	nd	nd
## 3	-20.180	0.280	-23.531	0.103	150.0	40	39/85	169.20	7.20	2.84	0.19	10.879
## 4	-18.770	0.200	-20.356	0.111	36.0	15	75	nd	nd	2.31	0.09	nd
## 5	-15.460	0.300	-17.254	0.118	10.0	2	91/97a	22.00	5.00	0.57	0.08	8.408
## 6	-15.400	0.200	-17.489	0.112	8.0	1	30/97a	19.90	2.40	0.35	0.04	8.105
##	V22	V23	V24	V25	V26	V27						
## 1	0.116	7.66	0.03	6.61	0.04	0.04						
## 2	nd	5.14	0.00	nd	nd	nd						
## 3	0.037	7.57	0.10	nd	nd	nd						
## 4	nd	6.87	0.15	nd	nd	nd						
## 5	0.130	6.16	0.11	nd	nd	nd						
## 6	0.069	6.07	0.05	nd	nd	nd						

Question 4: Large-scale climatic index (MEI): <https://www.esrl.noaa.gov/psd/enso/mei/data/meiv2.data>

Data set unavailable

```
# Data set no longer available
# url_mei <- "https://www.esrl.noaa.gov/psd/enso/mei/data/meiv2.data"

# df4 <- read.table(url_mei, sep = " ", header = FALSE, skip = 1, fill = FALSE)
# View(df4)
```

Question 5: Voice Data from Singing the Vowel 'ooh' <http://www.statsci.org/data/general/ooh.txt>

```
url_ooh <- "http://www.statsci.org/data/general/ooh.txt"

df5 <- read.table(url_ooh, header = TRUE)
str(df5)
```

```
## 'data.frame': 112 obs. of 1 variable:
## $ Magnitude: int 5 27 46 64 80 91 97 97 93 84 ...
```

```
View(df5)
```

Question 6: Air quality data: air2.dat

```
df6 <- read.table("air2.dat", sep = ";")
str(df6)

## 'data.frame': 153 obs. of 6 variables:
## $ Ozone : int 41 36 12 18 NA 28 23 19 8 NA ...
## $ Solar.R: int 190 118 149 313 NA NA 299 99 19 194 ...
## $ Wind : num 7.4 8 12.6 11.5 14.3 14.9 8.6 13.8 20.1 8.6 ...
## $ Temp : int 67 72 74 62 56 66 65 59 61 69 ...
## $ Month : int 5 5 5 5 5 5 5 5 5 5 ...
## $ Day : int 1 2 3 4 5 6 7 8 9 10 ...
```

Question 7: 2D Pollen counts: 8.Red-cross-feb-aug-89-starts 20 Feb-1989.xls

```
# 7. 2D Pollen counts: 8.Red-cross-feb-aug-89-starts 20 Feb-1989.xls (on Vula)
df7 <- read_excel("8.Red-cross-feb-aug-89-starts 20 Feb-1989.xls",
                  skip = 2, sheet = "2 D Pollen Counts")
df7$`Site: Observatory` <- as.factor(df7$`Site: Observatory`)
df7$Category <- as.factor(df7$Category)
str(df7)

## tibble [77 x 186] (S3: tbl_df/tbl/data.frame)
## $ Site: Observatory: Factor w/ 1 level "Red Cross Hospital Rondebosch": 1 1 1 1 1 1 1 1 1 1 ...
## $ Category : Factor w/ 5 levels "Fungal Spores",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ Classification : chr [1:77] "Alternaria" "Ascospores" "Aspergillus / Penicillium" "Basidiospores" ...
## $ Common Name : chr [1:77] NA NA NA NA ...
## $ 32559 : num [1:77] 51 32 0 6 0 0 3 3 64 0 ...
## $ 32560 : num [1:77] 13 19 0 26 0 0 13 0 3 0 ...
## $ 32561 : num [1:77] 157 205 0 3 0 0 29 10 250 0 ...
## $ 32562 : num [1:77] 35 35 0 157 0 0 19 0 13 0 ...
## $ 32563 : num [1:77] 269 26 0 3 0 0 16 26 218 0 ...
## $ 32564 : num [1:77] 0 6 0 3 0 0 0 0 3 0 ...
## $ 32565 : num [1:77] 0 16 0 0 0 0 0 0 51 0 ...
## $ 32566 : num [1:77] 61 2 0 0 0 0 6 13 99 0 ...
## $ 32567 : num [1:77] 342 131 0 3 0 0 544 10 1000 0 ...
## $ 32568 : num [1:77] 54 19 0 58 0 0 80 3 80 0 ...
## $ 32569 : num [1:77] 240 19 0 51 0 0 163 0 170 0 ...
## $ 32570 : num [1:77] 186 35 0 128 0 0 64 3 289 0 ...
## $ 32571 : num [1:77] 86 38 0 45 0 0 7 6 64 0 ...
## $ 32572 : num [1:77] 51 32 0 70 0 0 22 6 19 0 ...
## $ 32573 : num [1:77] 58 38 0 3 0 0 0 0 96 0 ...
## $ 32574 : num [1:77] 0 0 0 3 0 0 0 0 0 0 ...
## $ 32575 : num [1:77] 74 42 0 22 0 0 16 10 1000 0 ...
## $ 32576 : num [1:77] 58 42 0 3 0 0 35 6 205 0 ...
## $ 32577 : num [1:77] 77 10 0 19 0 0 26 0 64 0 ...
## $ 32578 : num [1:77] 16 13 0 10 0 0 58 0 13 0 ...
## $ 32579 : num [1:77] 48 26 0 6 0 0 48 3 67 0 ...
## $ 32580 : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32581 : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32582 : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32583 : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32584 : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32585 : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
```

```

## $ 32586      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32587      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32588      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32589      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32590      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32591      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32592      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32593      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32594      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32595      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32596      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32597      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32598      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32599      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32600      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32601      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32602      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32603      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32604      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32605      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32606      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32607      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32608      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32609      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32610      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32611      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32612      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32613      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32614      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32615      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32616      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32617      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32618      : num [1:77] 51 29 0 176 0 0 0 0 29 0 ...
## $ 32619      : num [1:77] 10 3 0 80 0 0 0 0 10 0 ...
## $ 32620      : num [1:77] 0 45 0 10 0 0 0 0 6 0 ...
## $ 32621      : num [1:77] 0 0 0 3 0 0 0 0 0 0 ...
## $ 32622      : num [1:77] 67 29 0 16 0 0 10 6 42 0 ...
## $ 32623      : num [1:77] 77 51 0 70 0 0 10 16 99 0 ...
## $ 32624      : num [1:77] 70 205 0 218 0 58 0 38 0 0 ...
## $ 32625      : num [1:77] 83 13 0 54 0 0 3 3 86 0 ...
## $ 32626      : num [1:77] 42 6 0 48 0 0 22 0 32 0 ...
## $ 32627      : num [1:77] 90 64 0 358 0 0 13 0 51 0 ...
## $ 32628      : num [1:77] 80 19 0 202 0 0 10 0 83 0 ...
## $ 32629      : num [1:77] 38 16 0 1000 0 0 22 0 64 0 ...
## $ 32630      : num [1:77] 22 38 0 1000 0 0 0 6 32 0 ...
## $ 32631      : num [1:77] 13 22 0 19 0 0 22 0 35 0 ...
## $ 32632      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32633      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32634      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32635      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32636      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32637      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32638      : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32639      : num [1:77] 42 3 0 32 0 0 3 10 45 0 ...

```

```
## $ 32640 : num [1:77] 10 100 0 32 0 0 0 3 26 0 ...
## $ 32641 : num [1:77] 0 109 0 32 0 0 0 0 0 0 ...
## $ 32642 : num [1:77] 0 10 0 48 0 0 3 0 0 0 ...
## $ 32643 : num [1:77] 0 29 32 0 0 0 0 0 0 0 ...
## $ 32644 : num [1:77] 0 3 0 0 0 0 6 0 0 0 ...
## $ 32645 : num [1:77] 13 13 0 0 0 0 0 3 13 0 ...
## $ 32646 : num [1:77] 29 90 0 22 0 0 3 0 32 0 ...
## $ 32647 : num [1:77] 26 32 0 112 0 0 10 0 3 0 ...
## $ 32648 : num [1:77] 1 22 0 141 0 0 3 0 16 0 ...
## $ 32649 : num [1:77] 3 26 0 835 0 0 0 0 3 0 ...
## $ 32650 : num [1:77] 32 54 0 109 0 0 19 0 13 0 ...
## $ 32651 : num [1:77] 29 32 0 880 0 0 22 0 29 0 ...
## $ 32652 : num [1:77] 0 0 0 0 0 0 0 0 0 0 ...
## $ 32653 : num [1:77] 10 19 0 99 0 0 3 0 26 0 ...
## [list output truncated]
```

Question 8: wader counts.xls

```
# 8. wader counts.xls (on Vula)
```

```
df8 <- read_excel("wader counts.XLS")
str(df8)
```

```
## tibble [155 x 67] (S3: tbl_df/tbl/data.frame)
## $ Site : chr [1:155] "Sandwich Harbour" "Sandwich Harbour" "Sandwich Harbour" "Sandwich Harbour" ...
## $ Robert's : num [1:155] 3 6 7 8 49 53 55 56 57 59 ...
## $ Species recorded: chr [1:155] "African Penguin (Spheniscus demersus)" "Great Crested Grebe (Podiceps ...
## $ 1977 01 07 : num [1:155] NA 50 713 6 308 3 NA NA NA NA ...
## $ 1978 01 10 : num [1:155] NA NA NA NA NA NA NA NA NA NA ...
## $ 1990 07 07 : num [1:155] NA NA NA 4 50 NA 8 NA NA NA ...
## $ 1991 02 11 : num [1:155] NA 2 NA 5 171 2 100 NA NA NA ...
## $ 1991 07 21 : num [1:155] NA 2 10 4 181 NA 73 142 NA NA ...
## $ 1991 09 11 : num [1:155] NA 3 8 NA 40 NA 111 320 NA NA ...
## $ 1992 01 19 : num [1:155] NA 1 2 1 520 NA 59 250 NA NA ...
## $ 1992 06 29 : num [1:155] NA NA 12 6 158 ...
## $ 1992 09 15 : num [1:155] NA 3 16 7 114 NA 56 132 NA NA ...
## $ 1992 11 18 : num [1:155] NA NA NA NA NA NA NA NA NA NA ...
## $ 1993 02 02 : num [1:155] NA 2 3 2 213 ...
## $ 1993 04 14 : num [1:155] NA 2 NA 2 246 1 91 4930 NA NA ...
## $ 1994 01 25 : num [1:155] NA NA NA 3 62 ...
## $ 1994 07 27 : num [1:155] NA 3 77 102 50 6 75 424 NA NA ...
## $ 1995 01 10 : num [1:155] NA 4 NA 4 48 ...
## $ 1995 04 27 : num [1:155] NA NA 10 1 102 4 46 504 NA NA ...
## $ 1995 07 20 : num [1:155] NA NA 4 28 112 2 74 5600 13 NA ...
## $ 1996 02 04 : num [1:155] NA 4 18 NA 214 1 15 5000 NA NA ...
## $ 1996 04 18 : num [1:155] NA NA NA NA 165 ...
## $ 1996 07 18 : num [1:155] NA 6 NA 6 27 ...
## $ 1997 01 23 : num [1:155] NA NA NA 1 603 NA 73 740 NA NA ...
## $ 1997 04 21 : num [1:155] NA NA NA NA 329 ...
## $ 1997 07 17 : num [1:155] NA 1 NA 4 36 ...
## $ 1998 01 27 : num [1:155] NA 13 NA 5 20 2 78 250 NA NA ...
## $ 1998 04 18 : num [1:155] NA NA NA NA 111 ...
## $ 1998 07 08 : num [1:155] NA NA 87 NA 82 NA 120 3500 NA NA ...
## $ 1998 11 03 : num [1:155] NA 3 NA 22 50 NA 32 3000 NA NA ...
```

```
## $ 1999 01 14 : num [1:155] NA 2 8 4 37 ...
## $ 1999 02 23 : num [1:155] NA 2 NA NA 550 NA 35 NA NA NA ...
## $ 1999 03 30 : num [1:155] NA NA 5 NA 29 NA 56 400 NA NA ...
## $ 1999 07 26 : num [1:155] NA 2 60 3 164 NA 105 100000 NA NA ...
## $ 1999 10 14 : num [1:155] NA NA 74 5 108 NA 248 13800 NA NA ...
## $ 1999 12 19 : num [1:155] NA NA NA NA 61 ...
## $ 2000 01 22 : num [1:155] NA NA 6 4 226 NA 80 7510 NA NA ...
## $ 2000 03 30 : num [1:155] NA 2 1 1 239 ...
## $ 2000 07 25 : num [1:155] NA 2 27 NA 103 ...
## $ 2000 10 17 : num [1:155] NA 1 14 9 99 ...
## $ 2001 01 18 : num [1:155] NA 2 NA 18 195 ...
## $ 2001 04 29 : num [1:155] NA 2 9 22 231 NA 84 2360 NA NA ...
## $ 2001 07 23 : num [1:155] NA 2 86 1 160 ...
## $ 2002 01 23 : num [1:155] 1 8 NA 8 197 2 146 16200 NA NA ...
## $ 2002 04 23 : num [1:155] NA 1 NA 8 577 ...
## $ 2002 07 25 : num [1:155] NA 3 534 21 64 1 243 200 NA NA ...
## $ 2003 01 16 : num [1:155] NA 8 2 21 191 NA 51 14000 NA NA ...
## $ 2003 04 25 : num [1:155] NA 5 120 7 653 NA 66 40000 NA 1 ...
## $ 2003 08 06 : num [1:155] NA NA 602 7 177 ...
## $ 2004 01 27 : num [1:155] NA 8 3 NA 520 NA 310 60000 NA NA ...
## $ 2004 04 30 : num [1:155] NA 5 1 2 900 ...
## $ 2004 07 24 : num [1:155] NA 2 2 8 420 NA 52 3250 NA NA ...
## $ 2005 02 02 : num [1:155] NA 18 15 NA 514 NA 153 84100 NA NA ...
## $ 2005 04 26 : num [1:155] NA 1 NA 1 402 ...
## $ 2005 07 26 : num [1:155] NA 1 280 3 51 NA 108 124 NA NA ...
## $ 2006 01 29 : num [1:155] NA 3 NA 3 35 NA 78 55 NA NA ...
## $ 2006 07 22 : num [1:155] NA 3 64 13 413 NA 70 77 NA NA ...
## $ 2007 02 04 : num [1:155] NA 2 16 1 124 ...
## $ 2007 07 23 : num [1:155] NA 1 70 29 85 ...
## $ 2008 01 20 : num [1:155] NA NA 70 31 54 ...
## $ 2008 07 19 : num [1:155] NA 2 154 10 41 ...
## $ 2009 01 24 : num [1:155] NA NA 184 7 32 NA 102 714 NA NA ...
## $ 2009 07 25 : num [1:155] NA 1 NA 12 71 1 71 13000 NA NA ...
## $ 2010 02 13 : num [1:155] NA NA NA 8 175 ...
## $ 2010 07 27 : num [1:155] NA 1 NA 15 50 3 20 100000 NA NA ...
## $ 2011 01 22 : num [1:155] NA 1 7 NA 130 NA 33 115 NA NA ...
## $ 2011 07 17 : num [1:155] NA NA NA NA 5 ...
```

```
# format as factors
```

```
df8$Site <- as.factor(df8$Site)
```

```
str(df8)
```

```
## tibble [155 x 67] (S3: tbl_df/tbl/data.frame)
```

```
## $ Site : Factor w/ 3 levels "Sandwich Harbour",...: 1 1 1 1 1 1 1 1 1 1 ...
```

```
## $ Robert's : num [1:155] 3 6 7 8 49 53 55 56 57 59 ...
```

```
## $ Species recorded: chr [1:155] "African Penguin (Spheniscus demersus)" "Great Crested Grebe (Podiceps
```

```
## $ 1977 01 07 : num [1:155] NA 50 713 6 308 3 NA NA NA NA ...
```

```
## $ 1978 01 10 : num [1:155] NA NA NA NA NA NA NA NA NA NA ...
```

```
## $ 1990 07 07 : num [1:155] NA NA NA 4 50 NA 8 NA NA NA ...
```

```
## $ 1991 02 11 : num [1:155] NA 2 NA 5 171 2 100 NA NA NA ...
```

```
## $ 1991 07 21 : num [1:155] NA 2 10 4 181 NA 73 142 NA NA ...
```

```
## $ 1991 09 11 : num [1:155] NA 3 8 NA 40 NA 111 320 NA NA ...
```

```
## $ 1992 01 19 : num [1:155] NA 1 2 1 520 NA 59 250 NA NA ...
```

```
## $ 1992 06 29 : num [1:155] NA NA 12 6 158 ...
```

```

## $ 1992 09 15 : num [1:155] NA 3 16 7 114 NA 56 132 NA NA ...
## $ 1992 11 18 : num [1:155] NA NA NA NA NA NA NA NA NA NA ...
## $ 1993 02 02 : num [1:155] NA 2 3 2 213 ...
## $ 1993 04 14 : num [1:155] NA 2 NA 2 246 1 91 4930 NA NA ...
## $ 1994 01 25 : num [1:155] NA NA NA 3 62 ...
## $ 1994 07 27 : num [1:155] NA 3 77 102 50 6 75 424 NA NA ...
## $ 1995 01 10 : num [1:155] NA 4 NA 4 48 ...
## $ 1995 04 27 : num [1:155] NA NA 10 1 102 4 46 504 NA NA ...
## $ 1995 07 20 : num [1:155] NA NA 4 28 112 2 74 5600 13 NA ...
## $ 1996 02 04 : num [1:155] NA 4 18 NA 214 1 15 5000 NA NA ...
## $ 1996 04 18 : num [1:155] NA NA NA NA 165 ...
## $ 1996 07 18 : num [1:155] NA 6 NA 6 27 ...
## $ 1997 01 23 : num [1:155] NA NA NA 1 603 NA 73 740 NA NA ...
## $ 1997 04 21 : num [1:155] NA NA NA NA 329 ...
## $ 1997 07 17 : num [1:155] NA 1 NA 4 36 ...
## $ 1998 01 27 : num [1:155] NA 13 NA 5 20 2 78 250 NA NA ...
## $ 1998 04 18 : num [1:155] NA NA NA NA 111 ...
## $ 1998 07 08 : num [1:155] NA NA 87 NA 82 NA 120 3500 NA NA ...
## $ 1998 11 03 : num [1:155] NA 3 NA 22 50 NA 32 3000 NA NA ...
## $ 1999 01 14 : num [1:155] NA 2 8 4 37 ...
## $ 1999 02 23 : num [1:155] NA 2 NA NA 550 NA 35 NA NA NA ...
## $ 1999 03 30 : num [1:155] NA NA 5 NA 29 NA 56 400 NA NA ...
## $ 1999 07 26 : num [1:155] NA 2 60 3 164 NA 105 100000 NA NA ...
## $ 1999 10 14 : num [1:155] NA NA 74 5 108 NA 248 13800 NA NA ...
## $ 1999 12 19 : num [1:155] NA NA NA NA 61 ...
## $ 2000 01 22 : num [1:155] NA NA 6 4 226 NA 80 7510 NA NA ...
## $ 2000 03 30 : num [1:155] NA 2 1 1 239 ...
## $ 2000 07 25 : num [1:155] NA 2 27 NA 103 ...
## $ 2000 10 17 : num [1:155] NA 1 14 9 99 ...
## $ 2001 01 18 : num [1:155] NA 2 NA 18 195 ...
## $ 2001 04 29 : num [1:155] NA 2 9 22 231 NA 84 2360 NA NA ...
## $ 2001 07 23 : num [1:155] NA 2 86 1 160 ...
## $ 2002 01 23 : num [1:155] 1 8 NA 8 197 2 146 16200 NA NA ...
## $ 2002 04 23 : num [1:155] NA 1 NA 8 577 ...
## $ 2002 07 25 : num [1:155] NA 3 534 21 64 1 243 200 NA NA ...
## $ 2003 01 16 : num [1:155] NA 8 2 21 191 NA 51 14000 NA NA ...
## $ 2003 04 25 : num [1:155] NA 5 120 7 653 NA 66 40000 NA 1 ...
## $ 2003 08 06 : num [1:155] NA NA 602 7 177 ...
## $ 2004 01 27 : num [1:155] NA 8 3 NA 520 NA 310 60000 NA NA ...
## $ 2004 04 30 : num [1:155] NA 5 1 2 900 ...
## $ 2004 07 24 : num [1:155] NA 2 2 8 420 NA 52 3250 NA NA ...
## $ 2005 02 02 : num [1:155] NA 18 15 NA 514 NA 153 84100 NA NA ...
## $ 2005 04 26 : num [1:155] NA 1 NA 1 402 ...
## $ 2005 07 26 : num [1:155] NA 1 280 3 51 NA 108 124 NA NA ...
## $ 2006 01 29 : num [1:155] NA 3 NA 3 35 NA 78 55 NA NA ...
## $ 2006 07 22 : num [1:155] NA 3 64 13 413 NA 70 77 NA NA ...
## $ 2007 02 04 : num [1:155] NA 2 16 1 124 ...
## $ 2007 07 23 : num [1:155] NA 1 70 29 85 ...
## $ 2008 01 20 : num [1:155] NA NA 70 31 54 ...
## $ 2008 07 19 : num [1:155] NA 2 154 10 41 ...
## $ 2009 01 24 : num [1:155] NA NA 184 7 32 NA 102 714 NA NA ...
## $ 2009 07 25 : num [1:155] NA 1 NA 12 71 1 71 13000 NA NA ...
## $ 2010 02 13 : num [1:155] NA NA NA 8 175 ...
## $ 2010 07 27 : num [1:155] NA 1 NA 15 50 3 20 100000 NA NA ...

```



```
## $ 2011 01 22      : num [1:155] NA 1 7 NA 130 NA 33 115 NA NA ...  
## $ 2011 07 17      : num [1:155] NA NA NA NA 5 ...
```

```
View\(df8\)
```