

# Time Series

*Alan Arnholt*

*8/30/2016*

Read in the data:

```
library(readr)
ages <- read_csv("Ch4Eg.csv")
ages
```

	Year	Women	Men
1	1998	25.0	26.7
2	1997	25.0	26.8
3	1996	24.8	27.1
4	1995	24.5	26.9
5	1994	24.5	26.7
6	1993	24.5	26.5
7	1992	24.4	26.5
8	1991	24.1	26.3
9	1990	23.9	26.1
10	1989	23.8	26.2
11	1988	23.6	25.9
12	1987	23.6	25.8
13	1986	23.1	25.7
14	1985	23.3	25.5
15	1984	23.0	25.4
16	1983	22.8	25.4
17	1982	22.5	25.2
18	1981	22.3	24.8
19	1980	22.0	24.7
20	1979	22.1	24.4
21	1978	21.8	24.2
22	1977	21.6	24.0
23	1976	21.3	23.8
24	1975	21.1	23.5
25	1974	21.1	23.1
26	1973	21.0	23.2
27	1972	20.9	23.3
28	1971	20.9	23.1
29	1970	20.8	23.2
30	1969	20.8	23.2
31	1968	20.8	23.1
32	1967	20.6	23.1
33	1966	20.5	22.8
34	1965	20.6	22.8
35	1964	20.5	23.1
36	1963	20.5	22.8
37	1962	20.3	22.7
38	1961	20.3	22.8
39	1960	20.3	22.8
40	1959	20.2	22.5

```

41 1958 20.2 22.6
42 1957 20.3 22.6
43 1956 20.1 22.5
44 1955 20.2 22.6
45 1954 20.3 23.0
46 1953 20.2 22.8
47 1952 20.2 23.0
48 1951 20.4 22.9
49 1950 20.3 22.8
50 1949 20.3 22.7
51 1948 20.4 23.3
52 1947 20.5 23.7
53 1940 21.5 24.3
54 1930 21.3 24.3
55 1920 21.2 24.6
56 1910 21.6 25.1
57 1900 21.9 25.9
58 1890 22.0 26.1

```

Create side-by-side boxplots:

```

library(tidyr)
NDF <- gather(ages, Gender, Age, -Year)
NDF

```

```

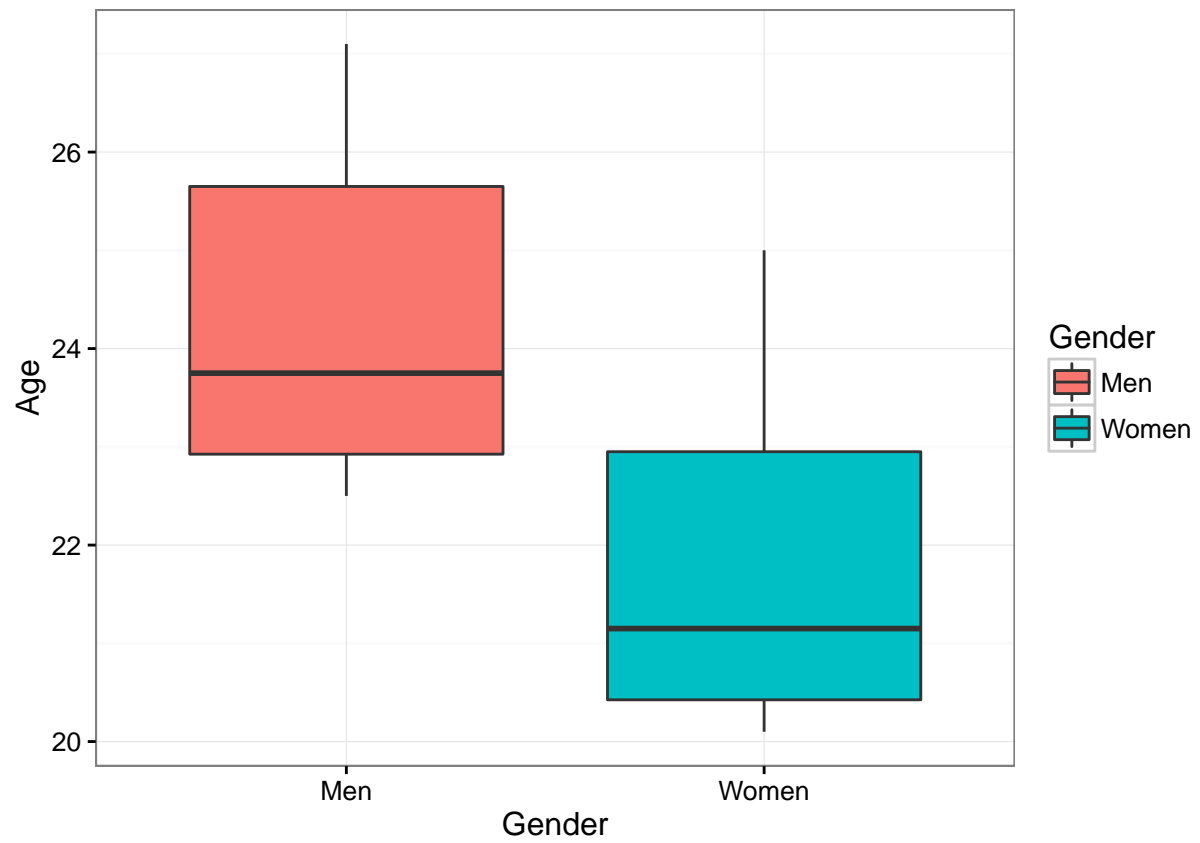
# A tibble: 116 x 3
   Year Gender   Age
   <int> <chr> <dbl>
1  1998 Women  25.0
2  1997 Women  25.0
3  1996 Women  24.8
4  1995 Women  24.5
5  1994 Women  24.5
6  1993 Women  24.5
7  1992 Women  24.4
8  1991 Women  24.1
9  1990 Women  23.9
10 1989 Women  23.8
# ... with 106 more rows

```

```

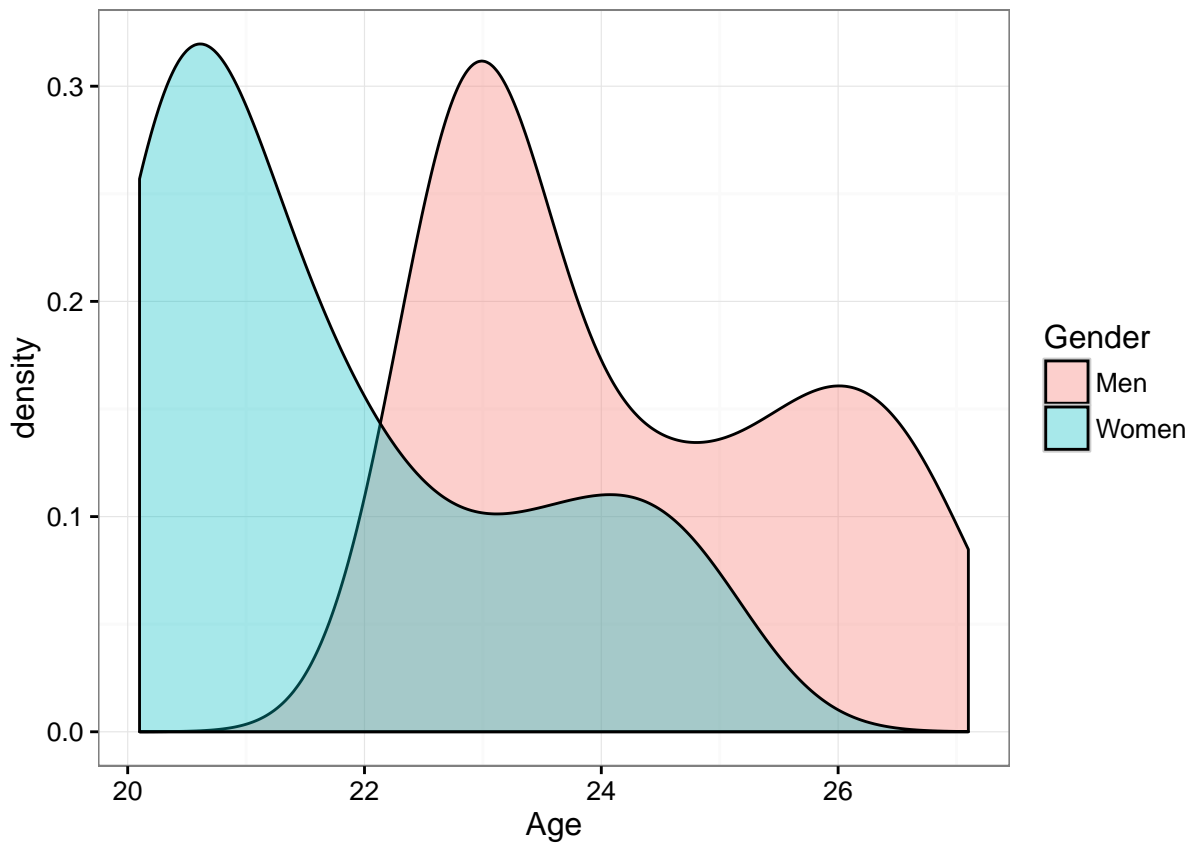
library(ggplot2)
ggplot(data = NDF, aes(x = Gender, y = Age, fill = Gender)) + geom_boxplot() +
  theme_bw()

```



Density plots:

```
ggplot(data = NDF, aes(x = Age, fill = Gender)) +  
  geom_density(alpha = 0.35) +  
  theme_bw()
```



Summary information:

```
library(dplyr)
SI <- NDF %>%
  group_by(Gender) %>%
  summarise(av_age = mean(Age), md_age = median(Age), sd_age = sd(Age))
SI
```

```
# A tibble: 2 x 4
  Gender av_age md_age sd_age
  <chr>   <dbl> <dbl> <dbl>
1 Men    24.25000 23.75 1.490820
2 Women  21.75172 21.15 1.554421
```

Create a time-series plot:

```
ggplot(data = ages, aes(x = Year, y = Women)) +
  geom_line(color = "purple") +
  geom_point(color = "purple") +
  geom_smooth(color = "pink") +
  geom_line(aes(x = Year, y = Men), color = "blue") +
  geom_point(aes(x = Year, y = Men), color = "blue") +
  geom_smooth(aes(x = Year, y = Men), color = "lightblue") +
  labs(y = "Age") +
  theme_bw()
```

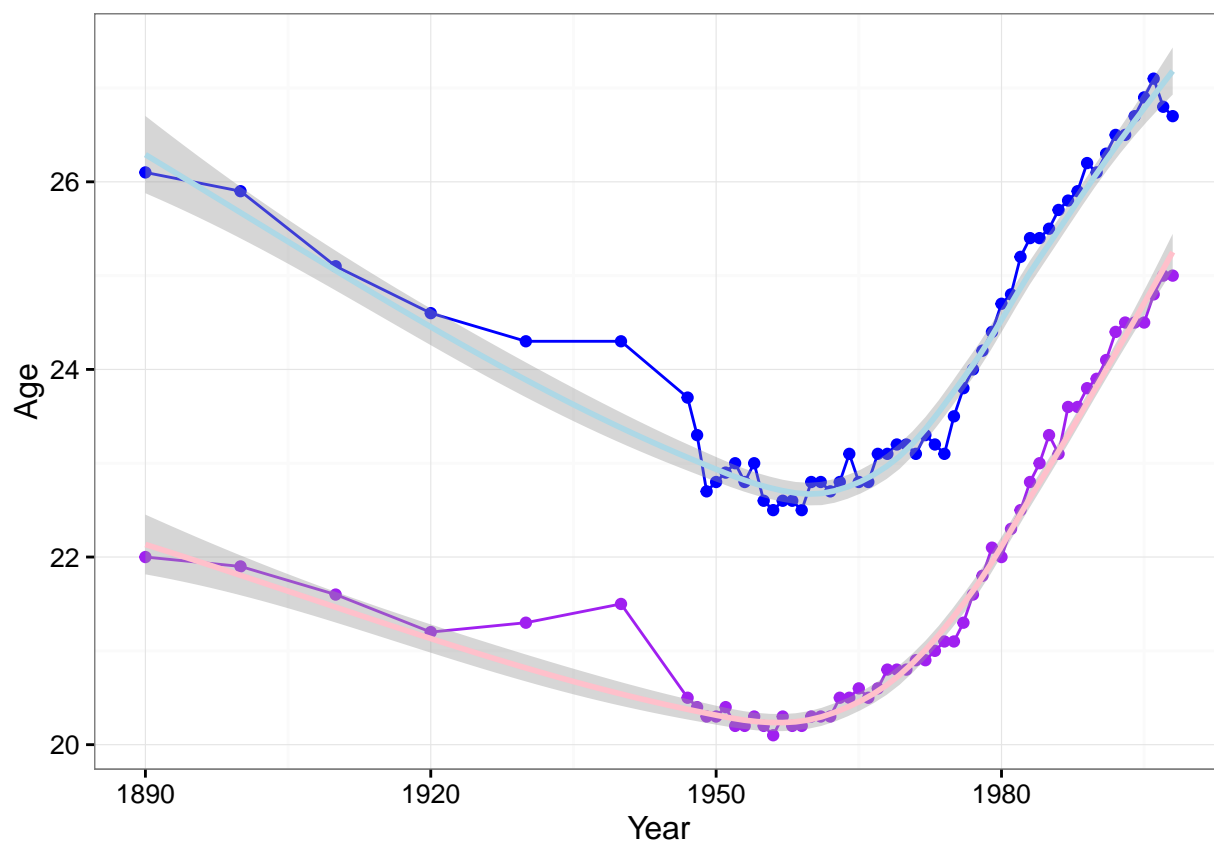


Figure 1: Time Series Plot

Another one using NDF:

```
ggplot(data = NDF, aes(x = Year, y = Age, color = Gender)) +  
  geom_point() +  
  geom_line() +  
  geom_smooth() +  
  theme_bw()
```

