Assignment 1

Erling Tennøy Nordtvedt Oleg Karpov

2024-09-16

Excercise 1 - R syntax & data structures

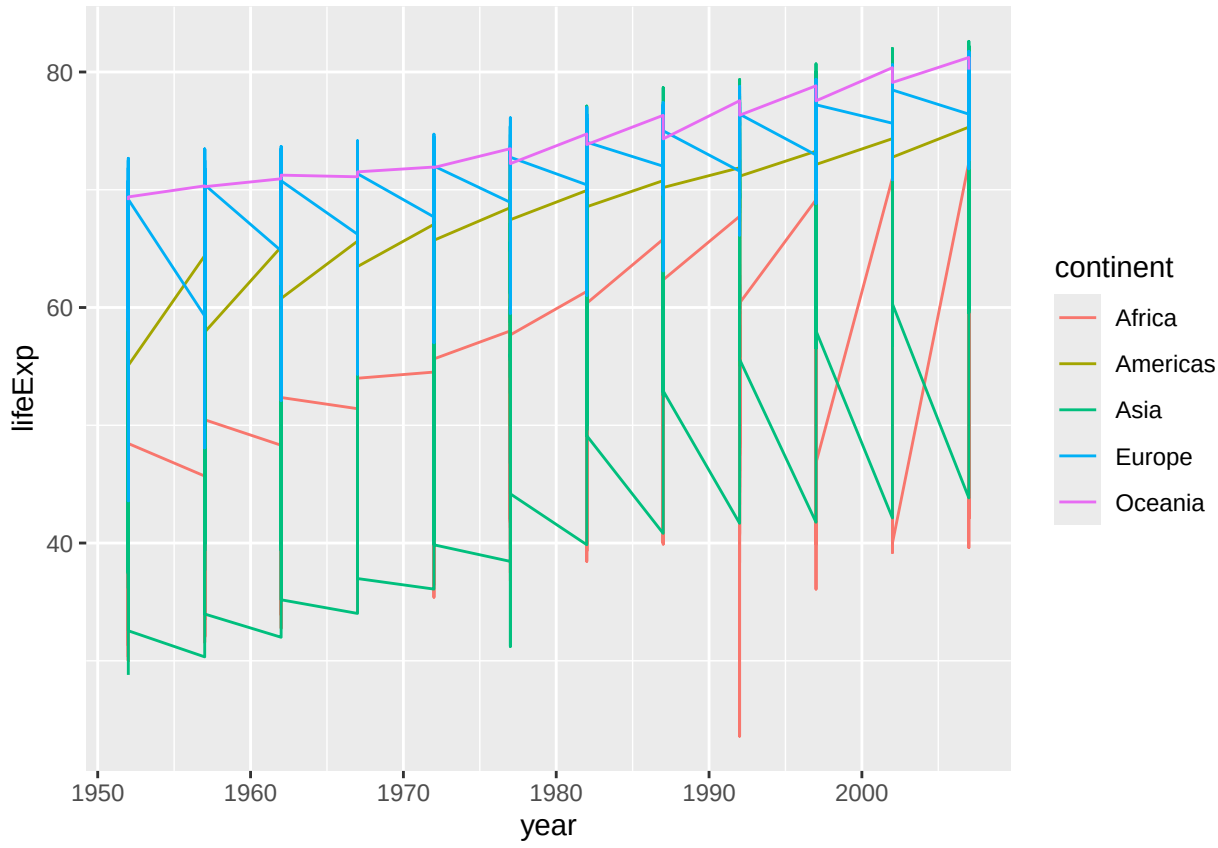
(a)

```
gapminder <- read.csv("gapminder.csv")
summary(gapminder)
```

```
##        X              country           continent              year     
##  Min.   :   1.0   Length:1704        Length:1704        Min.   :1952  
##  1st Qu.: 426.8   Class :character   Class :character   1st Qu.:1966  
##  Median : 852.5   Mode  :character   Mode  :character   Median :1980  
##  Mean   : 852.5                                        Mean   :1980  
##  3rd Qu.:1278.2                                        3rd Qu.:1993  
##  Max.   :1704.0                                        Max.   :2007  
##     lifeExp           pop             gdpPercap       
##  Min.   :23.60   Min.   :6.001e+04   Min.   :   241.2  
##  1st Qu.:48.20   1st Qu.:2.794e+06   1st Qu.:  1202.1  
##  Median :60.71   Median :7.024e+06   Median :  3531.8  
##  Mean   :59.47   Mean   :2.960e+07   Mean   :  7215.3  
##  3rd Qu.:70.85   3rd Qu.:1.959e+07   3rd Qu.:  9325.5  
##  Max.   :82.60   Max.   :1.319e+09   Max.   :113523.1
```

(b)

```
gapminder %>%
  ggplot(aes(x=year, y=lifeExp, colour=continent)) +
    #geom_bar(position='dodge', stat='identity')
    geom_line()
```



(c)

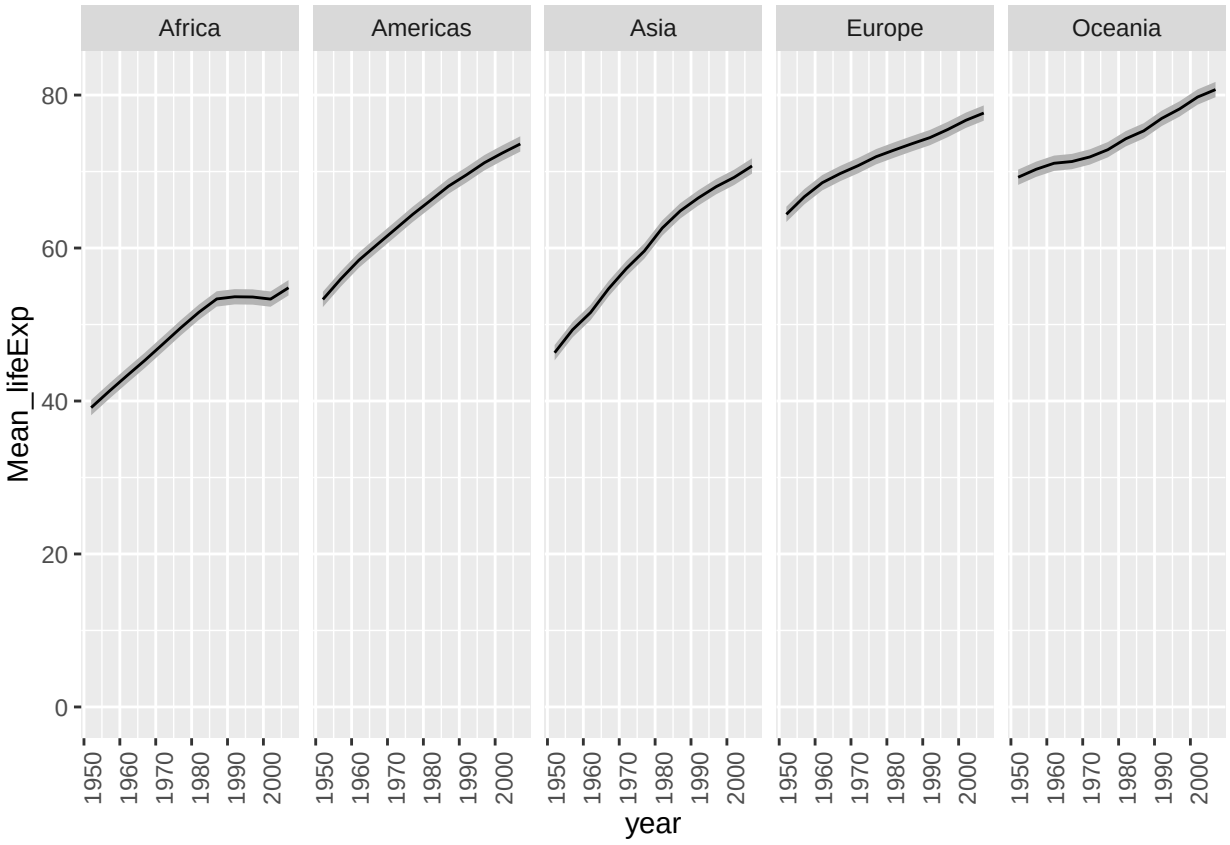
```
gapminder %>%
  group_by(continent, year) %>%
  summarise_at(vars(lifeExp), list(Min = min, Med = median, Mean = mean, Max = max, Sd = sd)) %>%
  data.frame()
```

```
##     continent year    Min     Med     Mean    Max          Sd
## 1      Africa 1952 30.000 38.8330 39.13550 52.724  5.15158143
## 2      Africa 1957 31.570 40.5925 41.26635 58.089  5.62012285
## 3      Africa 1962 32.767 42.6305 43.31944 60.246  5.87536393
## 4      Africa 1967 34.113 44.6985 45.33454 61.557  6.08267263
## 5      Africa 1972 35.400 47.0315 47.45094 64.274  6.41625832
## 6      Africa 1977 36.788 49.2725 49.58042 67.064  6.80819741
## 7      Africa 1982 38.445 50.7560 51.59287 69.885  7.37594009
## 8      Africa 1987 39.906 51.6395 53.34479 71.913  7.86408911
## 9      Africa 1992 23.599 52.4290 53.62958 73.615  9.46107099
## 10     Africa 1997 36.087 52.7590 53.59827 74.772  9.10338658
## 11     Africa 2002 39.193 51.2355 53.32523 75.744  9.58649585
## 12     Africa 2007 39.613 52.9265 54.80604 76.442  9.63078067
## 13   Americas 1952 37.579 54.7450 53.27984 68.750  9.32608188
## 14   Americas 1957 40.696 56.0740 55.96028 69.960  9.03319228
## 15   Americas 1962 43.428 58.2990 58.39876 71.300  8.50354374
## 16   Americas 1967 45.032 60.5230 60.41092 72.130  7.90917104
## 17   Americas 1972 46.714 63.4410 62.39492 72.880  7.32301680
## 18   Americas 1977 49.923 66.3530 64.39156 74.210  7.06949562
```

```
## 19  Americas 1982 51.461 67.4050 66.22884 75.760  6.72083382
## 20  Americas 1987 53.636 69.4980 68.09072 76.860  5.80192884
## 21  Americas 1992 55.089 69.8620 69.56836 77.950  5.16710381
## 22  Americas 1997 56.671 72.1460 71.15048 78.610  4.88758390
## 23  Americas 2002 58.137 72.0470 72.42204 79.770  4.79970550
## 24  Americas 2007 60.916 72.8990 73.60812 80.653  4.44094763
## 25      Asia 1952 28.801 44.8690 46.31439 65.390  9.29175070
## 26      Asia 1957 30.332 48.2840 49.31854 67.840  9.63542862
## 27      Asia 1962 31.997 49.3250 51.56322 69.390  9.82063194
## 28      Asia 1967 34.020 53.6550 54.66364 71.430  9.65096458
## 29      Asia 1972 36.088 56.9500 57.31927 73.420  9.72270004
## 30      Asia 1977 31.220 60.7650 59.61056 75.380 10.02219698
## 31      Asia 1982 39.854 63.7390 62.61794 77.110  8.53522141
## 32      Asia 1987 40.822 66.2950 64.85118 78.670  8.20379188
## 33      Asia 1992 41.674 68.6900 66.53721 79.360  8.07554897
## 34      Asia 1997 41.763 70.2650 68.02052 80.690  8.09117061
## 35      Asia 2002 42.129 71.0280 69.23388 82.000  8.37459539
## 36      Asia 2007 43.828 72.3960 70.72848 82.603  7.96372447
## 37    Europe 1952 43.585 65.9000 64.40850 72.670  6.36108825
## 38    Europe 1957 48.079 67.6500 66.70307 73.470  5.29580539
## 39    Europe 1962 52.098 69.5250 68.53923 73.680  4.30249956
## 40    Europe 1967 54.336 70.6100 69.73760 74.160  3.79972850
## 41    Europe 1972 57.005 70.8850 70.77503 74.720  3.24057637
## 42    Europe 1977 59.507 72.3350 71.93777 76.110  3.12102998
## 43    Europe 1982 61.036 73.4900 72.80640 76.990  3.21826030
## 44    Europe 1987 63.108 74.8150 73.64217 77.410  3.16968034
## 45    Europe 1992 66.146 75.4510 74.44010 78.770  3.20978109
## 46    Europe 1997 68.835 76.1160 75.50517 79.390  3.10467655
## 47    Europe 2002 70.845 77.5365 76.70060 80.620  2.92217958
## 48    Europe 2007 71.777 78.6085 77.64860 81.757  2.97981266
## 49   Oceania 1952 69.120 69.2550 69.25500 69.390  0.19091883
## 50   Oceania 1957 70.260 70.2950 70.29500 70.330  0.04949747
## 51   Oceania 1962 70.930 71.0850 71.08500 71.240  0.21920310
## 52   Oceania 1967 71.100 71.3100 71.31000 71.520  0.29698485
## 53   Oceania 1972 71.890 71.9100 71.91000 71.930  0.02828427
## 54   Oceania 1977 72.220 72.8550 72.85500 73.490  0.89802561
## 55   Oceania 1982 73.840 74.2900 74.29000 74.740  0.63639610
## 56   Oceania 1987 74.320 75.3200 75.32000 76.320  1.41421356
## 57   Oceania 1992 76.330 76.9450 76.94500 77.560  0.86974134
## 58   Oceania 1997 77.550 78.1900 78.19000 78.830  0.90509668
## 59   Oceania 2002 79.110 79.7400 79.74000 80.370  0.89095454
## 60   Oceania 2007 80.204 80.7195 80.71950 81.235  0.72902709
```

(d)

```
gapminder %>%
  group_by(continent, year) %>%
  summarise(Mean_lifeExp= mean(lifeExp, na.rm = T), .groups = 'drop') %>%
  ggplot(aes(x=year, y=Mean_lifeExp)) +
  geom_ribbon(aes(ymin= Mean_lifeExp - 1, ymax = Mean_lifeExp + 1), fill = "grey70") +
  geom_line() +
  facet_grid(.~continent) +
  theme(axis.text.x = element_text(angle=90)) +
  ylim(0, NA)
```



Excercise 2 - Elementary data analysis and model training

(a)

```
#weatherHistory <- read.csv("weatherHistory.csv")
#head(weatherHistory)
#weatherHistory %>%
#  ggplot(aes(x=Formatted.Date, y=Temperature..C.))
```