assignemnt1_report

Nina Dobša

2024-04-23

Assignment 1 - Weather forecast analysis for Nairobi

Libraries

```
library(rvest)
library(stringr)
library(ggplot2)
library(knitr)
library(httr)
```

During coding I used some usual libraries for web-scraping (rvest), string manipulation (stringr), plotting (ggplot2), table displaying (knitr) and specific library for making HTTP requests (httr).

Functions used

For web scraping tasks, I used several functions from the rvest library. Initially, I employed the *read_html()* function to retrieve the HTML content of the web page. Then, to extract specific sections of the HTML document, I used the *html_elements()* and *html_nodes()* functions. Once the relevant sections were identified, I employed the *html_table()* function to extract tabular data from sections or *html_text()* to extract text data. During scraping I encountered some language-related issues. To solve them, I utilized the GET() function from the httr library to specify the desired language settings.

Below you can find example of web-scraping for the a) part of the task.

```
response <- GET("https://www.timeanddate.com/weather/kenya/nairobi",
                add_headers(`Accept-Language` = "en-US,en;q=0.9"))
html_content <- content(response, as = "text")

# Parse the HTML content
html_w48 <- read_html(html_content, encoding = "UTF-8")
# getting a table from html document
table_html <- html_w48 %>% html_elements('#wt-48') %>% html_table(header = NA)
```

For displaying tables I used *kable()* function which you will see later on. For graphing, library ggplot2 helped me in getting well shaped graphs similar to the graphs on the website.

a) Table for the next 48 hours forecast

For the fist part of the task, I had to retrieve the data from the summary table „Forecast for the next 48 hours“.

Data from the website:

Forecast for the next 48 hours

	Wednesday		Thursday				Friday
	Afternoon	Evening	Night	Morning	Afternoon	Evening	Night
Forecast							
Temperature	22 °C	19 °C	17 °C	19 °C	21 °C	18 °C	16 °C
	A few thunder-storms. Overcast.	Thunderstorms. Overcast.	Rain showers. Overcast.	Passing showers. Overcast.	Passing showers. Overcast.	Isolated thunder-storms. Overcast.	Showery. Overcast.
Feels Like	23 °C	19 °C	17 °C	19 °C	21 °C	18 °C	16 °C
Wind Speed	12 km/h	6 km/h	3 km/h	4 km/h	2 km/h	8 km/h	4 km/h
Wind Direction	SSE ↖	E ←	SE ↖	W →	SW ↗	ENE ↙	N ↓
Humidity	71%	86%	93%	84%	71%	84%	93%
Dew Point	16 °C	16 °C	16 °C	16 °C	15 °C	15 °C	15 °C
Visibility	6 km	3 km	5 km	6 km	12 km	7 km	3 km
Probability of Precipitation	53%	56%	44%	46%	56%	47%	48%
Amount of Rain	3.5 mm	10.4 mm	3.6 mm	3.0 mm	3.7 mm	3.0 mm	14.8 mm

* Updated Wednesday, 24 April 2024 09:48:10 Nairobi time - Weather by CustomWeather, © 2024

Table for the next 48 hours forecast:

```
# extracting the first (and the only) element of list table_html to get the table
table_w48 <- table_html[[1]]

# deleting row Forecast and the last row
table_w48 <- table_w48[-c(2,13), ]

# printing the table
kable(table_w48)
```

	Wednesday	Wednesday	Thursday	Thursday	Thursday	Thursday	Friday
	Afternoon	Evening	Night	Morning	Afternoon	Evening	Night
Temperature	22 °C	19 °C	17 °C	19 °C	21 °C	18 °C	16 °C
	A few thunder-storms. Overcast.	Thunderstorms. Overcast.	Rain showers. Overcast.	Passing showers. Overcast.	Passing showers. Overcast.	Isolated thunder-storms. Overcast.	Showery. Overcast.
Feels Like	23 °C	19 °C	17 °C	19 °C	21 °C	18 °C	16 °C
Wind Speed	12 km/h	6 km/h	3 km/h	4 km/h	2 km/h	8 km/h	4 km/h
Wind Direction	SSE↑	E↑	SE↑	W↑	SW↑	ENE↑	N↑
Humidity	71%	86%	93%	84%	71%	84%	93%
Dew Point	16 °C	16 °C	16 °C	16 °C	15 °C	15 °C	15 °C

	Wednesday	Wednesday	Thursday	Thursday	Thursday	Thursday	Friday
Visibility	6 km	3 km	5 km	6 km	12 km	7 km	3 km
Probability of Precipitation	53%	56%	44%	46%	56%	47%	48%
Amount of Rain	3.5 mm	10.4 mm	3.6 mm	3.0 mm	3.7 mm	3.0 mm	14.8 mm

b) Table and graph for the next 24 hours forecast

Data from the website:

	Conditions		Comfort			Precipitation	
Time	Temp	Weather	Feels Like	Wind	Humidity	Chance	Amount
13:00 Wed, 24 Apr	21 °C	A few tstorms. Overcast.	22 °C	12 km/h	75%	54%	0.7 mm (rain)
14:00	22 °C	A few tstorms. Overcast.	23 °C	12 km/h	73%	54%	0.6 mm (rain)
15:00	22 °C	A few tstorms. Overcast.	23 °C	12 km/h	71%	53%	0.6 mm (rain)
16:00	22 °C	A few tstorms. Overcast.	23 °C	11 km/h	73%	52%	0.5 mm (rain)
17:00	21 °C	A few tstorms. Overcast.	21 °C	10 km/h	75%	51%	0.5 mm (rain)
18:00	21 °C	A few tstorms. Overcast.	21 °C	9 km/h	77%	52%	0.6 mm (rain)
19:00	20 °C	Thunderstorms. Overcast.	20 °C	8 km/h	80%	54%	1.2 mm (rain)
20:00	19 °C	Rain showers. Overcast.	19 °C	7 km/h	83%	55%	2.2 mm (rain)
21:00	19 °C	Rain showers. Overcast.	19 °C	6 km/h	86%	56%	2.7 mm (rain)
22:00	18 °C	Rain showers. Overcast.	18 °C	5 km/h	87%	53%	2.3 mm (rain)
23:00	18 °C	Light showers. Overcast.	18 °C	5 km/h	90%	49%	1.3 mm (rain)
Time	Temp	Weather	Feels Like	Wind	Humidity	Chance	Amount
00:00 Thu, 25 Apr	17 °C	Light showers. Overcast.	17 °C	5 km/h	91%	45%	0.7 mm (rain)
01:00	17 °C	A few showers. Overcast.	17 °C	4 km/h	91%	43%	0.7 mm (rain)
02:00	17 °C	A few showers. Overcast.	17 °C	2 km/h	92%	44%	0.6 mm (rain)
03:00	17 °C	A few showers. Overcast.	17 °C	3 km/h	93%	44%	0.6 mm (rain)
04:00	17 °C	A few showers. Overcast.	17 °C	3 km/h	93%	42%	0.6 mm (rain)
05:00	17 °C	A few showers. Overcast.	17 °C	4 km/h	94%	37%	0.6 mm (rain)
06:00	17 °C	A few showers. Overcast.	17 °C	5 km/h	94%	34%	0.5 mm (rain)
07:00	17 °C	A few showers. Overcast.	17 °C	4 km/h	94%	35%	0.6 mm (rain)
08:00	18 °C	Light showers. Overcast.	18 °C	4 km/h	90%	40%	0.7 mm (rain)
09:00	19 °C	Light showers. Overcast.	19 °C	4 km/h	84%	46%	0.8 mm (rain)
10:00	20 °C	A few showers. Overcast.	20 °C	5 km/h	78%	51%	0.6 mm (rain)
11:00	20 °C	Passing showers. Overcast.	20 °C	6 km/h	73%	53%	0.3 mm (rain)
12:00	21 °C	Passing showers. Overcast.	23 °C	8 km/h	70%	35%	0.1 mm (rain)

Table for the next 24 hours forecast:

Time	Temp	Weather	Feels Like	Wind	Humidity	Chance	Amount
14:00Wed, 24 Apr	22 °C	A few tstorms. Overcast.	23 °C	12 km/h ↑	73%	54%	0.6 mm (rain)
15:00	22 °C	A few tstorms. Overcast.	23 °C	12 km/h ↑	71%	53%	0.6 mm (rain)
16:00	22 °C	A few tstorms. Overcast.	23 °C	11 km/h ↑	73%	52%	0.5 mm (rain)
17:00	21 °C	A few tstorms. Overcast.	21 °C	10 km/h ↑	75%	51%	0.5 mm (rain)
18:00	21 °C	A few tstorms. Overcast.	21 °C	9 km/h ↑	77%	52%	0.6 mm (rain)
19:00	20 °C	Thunderstorms. Overcast.	20 °C	8 km/h ↑	80%	54%	1.2 mm (rain)
20:00	19 °C	Rain showers. Overcast.	19 °C	7 km/h ↑	83%	55%	2.2 mm (rain)
21:00	19 °C	Rain showers. Overcast.	19 °C	6 km/h ↑	86%	56%	2.7 mm (rain)
22:00	18 °C	Rain showers. Overcast.	18 °C	5 km/h ↑	87%	53%	2.3 mm (rain)
23:00	18 °C	Light showers. Overcast.	18 °C	5 km/h ↑	90%	49%	1.3 mm (rain)
00:00Thu, 25 Apr	17 °C	Light showers. Overcast.	17 °C	5 km/h ↑	91%	45%	0.7 mm (rain)
01:00	17 °C	A few showers. Overcast.	17 °C	4 km/h ↑	91%	43%	0.7 mm (rain)
02:00	17 °C	A few showers. Overcast.	17 °C	2 km/h ↑	92%	44%	0.6 mm (rain)
03:00	17 °C	A few showers. Overcast.	17 °C	3 km/h ↑	93%	44%	0.6 mm (rain)
04:00	17 °C	A few showers. Overcast.	17 °C	3 km/h ↑	93%	42%	0.6 mm (rain)
05:00	17 °C	A few showers. Overcast.	17 °C	4 km/h ↑	94%	37%	0.6 mm (rain)
06:00	17 °C	A few showers. Overcast.	17 °C	5 km/h ↑	94%	34%	0.5 mm (rain)
07:00	17 °C	A few showers. Overcast.	17 °C	4 km/h ↑	94%	35%	0.6 mm (rain)
08:00	18 °C	Light showers. Overcast.	18 °C	4 km/h ↑	90%	40%	0.7 mm (rain)

Time	Temp	Weather	Feels Like	Wind	Humidity	Chance	Amount
09:00	19 °C	Light showers. Overcast.	19 °C	4 km/h ↑	84%	46%	0.8 mm (rain)
10:00	20 °C	A few showers. Overcast.	20 °C	5 km/h ↑	78%	51%	0.6 mm (rain)
11:00	20 °C	Passing showers. Overcast.	20 °C	6 km/h ↑	73%	53%	0.3 mm (rain)
12:00	21 °C	Passing showers. Overcast.	23 °C	8 km/h ↑	70%	35%	0.1 mm (rain)
13:00	21 °C	Passing showers. Overcast.	23 °C	6 km/h ↑	69%	41%	0.1 mm (rain)

Since the task is to graph temperature, wind and amount values depending on time, I had to convert some values to numeric / datetime type. For columns Temp, Amount and Wind I firstly extracted numbers and then converted them to numeric with functions str_extract() and as.numeric(). For the time column, I decided to make two new columns: DateTime and Date. DateTime column is type POSIXct gotten from Time column using function as.POSIXct(), this column will be used for graphing, and the column Time will just be used for labeling x-axis. As POSIXct function converts column Time supposing that all hours are from today, I had to add 86 400s (24 hours) to the DateTime values after midnight. The column Date holds the information about a date.

Table for next 24 hours forecast after transformations:

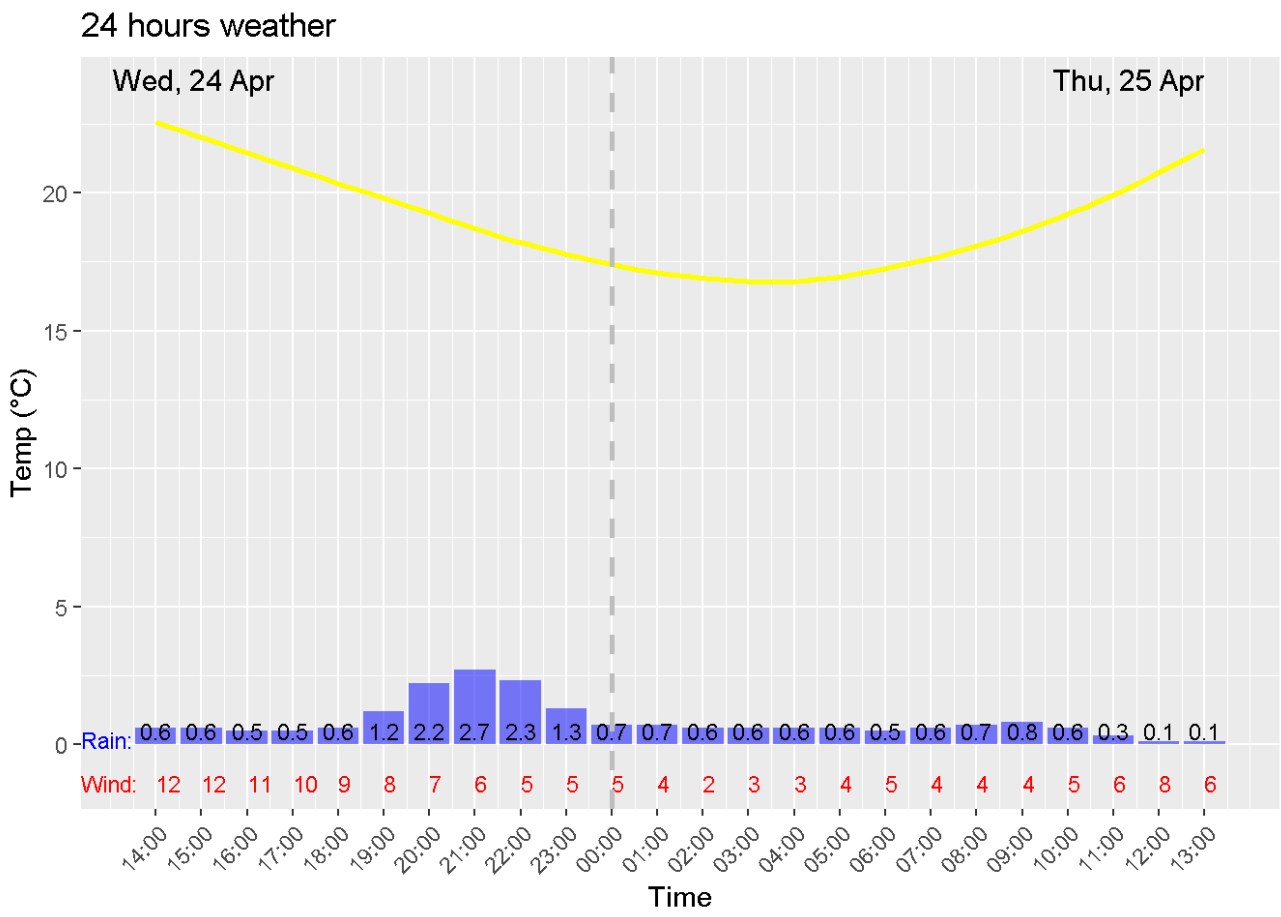
Time	Temp	Weather	Feels Like	Wind	Humidity	Chance	Amount	DateTime	Date
14:00	22	A few tstorms. Overcast.	23 °C	12 ↑	73%	54%	0.6	1713956400	Wed, 24 Apr
15:00	22	A few tstorms. Overcast.	23 °C	12 ↑	71%	53%	0.6	1713960000	Wed, 24 Apr
16:00	22	A few tstorms. Overcast.	23 °C	11 ↑	73%	52%	0.5	1713963600	Wed, 24 Apr
17:00	21	A few tstorms. Overcast.	21 °C	10 ↑	75%	51%	0.5	1713967200	Wed, 24 Apr
18:00	21	A few tstorms. Overcast.	21 °C	9 ↑	77%	52%	0.6	1713970800	Wed, 24 Apr
19:00	20	Thunderstorms. Overcast.	20 °C	8 ↑	80%	54%	1.2	1713974400	Wed, 24 Apr

Time	Temp	Weather	Feels Like	Wind	Humidity	Chance	Amount	DateTime	Date
20:00	19	Rain showers. Overcast.	19 °C	7 ↑	83%	55%	2.2	1713978000	Wed, 24 Apr
21:00	19	Rain showers. Overcast.	19 °C	6 ↑	86%	56%	2.7	1713981600	Wed, 24 Apr
22:00	18	Rain showers. Overcast.	18 °C	5 ↑	87%	53%	2.3	1713985200	Wed, 24 Apr
23:00	18	Light showers. Overcast.	18 °C	5 ↑	90%	49%	1.3	1713988800	Wed, 24 Apr
00:00	17	Light showers. Overcast.	17 °C	5 ↑	91%	45%	0.7	1713992400	Thu, 25 Apr
01:00	17	A few showers. Overcast.	17 °C	4 ↑	91%	43%	0.7	1713996000	Thu, 25 Apr
02:00	17	A few showers. Overcast.	17 °C	2 ↑	92%	44%	0.6	1713999600	Thu, 25 Apr
03:00	17	A few showers. Overcast.	17 °C	3 ↑	93%	44%	0.6	1714003200	Thu, 25 Apr
04:00	17	A few showers. Overcast.	17 °C	3 ↑	93%	42%	0.6	1714006800	Thu, 25 Apr
05:00	17	A few showers. Overcast.	17 °C	4 ↑	94%	37%	0.6	1714010400	Thu, 25 Apr
06:00	17	A few showers. Overcast.	17 °C	5 ↑	94%	34%	0.5	1714014000	Thu, 25 Apr
07:00	17	A few showers. Overcast.	17 °C	4 ↑	94%	35%	0.6	1714017600	Thu, 25 Apr
08:00	18	Light showers. Overcast.	18 °C	4 ↑	90%	40%	0.7	1714021200	Thu, 25 Apr
09:00	19	Light showers. Overcast.	19 °C	4 ↑	84%	46%	0.8	1714024800	Thu, 25 Apr

Time	Temp	Weather	Feels Like	Wind	Humidity	Chance	Amount	DateTime	Date
10:00	20	A few showers. Overcast.	20 °C	5 ↑	78%	51%	0.6	1714028400	Thu, 25 Apr
11:00	20	Passing showers. Overcast.	20 °C	6 ↑	73%	53%	0.3	1714032000	Thu, 25 Apr
12:00	21	Passing showers. Overcast.	23 °C	8 ↑	70%	35%	0.1	1714035600	Thu, 25 Apr
13:00	21	Passing showers. Overcast.	23 °C	6 ↑	69%	41%	0.1	1714039200	Thu, 25 Apr

Graph for the next 24 hours forecast:

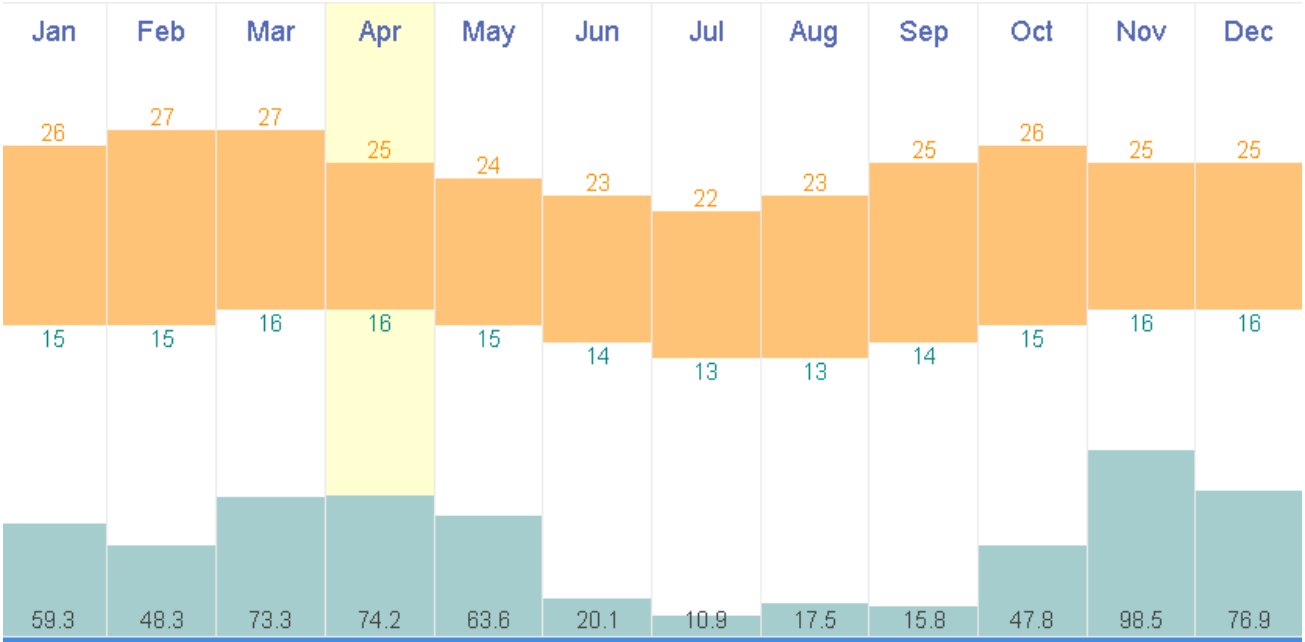
```
## `geom_smooth()` using formula = 'y ~ x'
```



c) Table, graph and avarage values for monthly annual weather

For the last part of the task, I retrieved the monthly annual weather averages from „Climate (Averages) - Annual Weather Averages“, calculated summary statistics (annual minimum, maximum, and mean temperature annual mean precipitation) and graphed the monthly values.

Data from the website:



April Climate & Weather Averages in Nairobi

High Temp: 25 °C	Precipitation: 74.2 mm	Wind: 21 km/h
Low Temp: 16 °C	Humidity: 74%	Pressure: 1015 mbar
Mean Temp: 21 °C	Dew Point: 15 °C	Visibility: 15 km

Web scraping for this part of the task was a bit tricky. I decided to scrape information for each month separably and then put the informations together in a table form. Below you can find an example of scraping for January.

```
# Read the HTML content from the website
html <- read_html("https://www.timeanddate.com/weather/kenya/nairobi/climate")

# Select all <p> elements within the specified path
january_info <- html %>%
  html_nodes("#climateTable > div.climate-month.climate-month--january > * > p") %>%
  html_text()
```

Table of monthly annual weather:

Month	High Temp	Low Temp	Mean Temp	Precipitation	Humidity	Dew Point	Wind	Pressure	Visibility
1	26	15	21	59.3	65	13	25	1014	15
2	27	15	21	48.3	60	12	26	1014	15

Month	High Temp	Low Temp	Mean Temp	Precipitation	Humidity	Dew Point	Wind	Pressure	Visibility
3	27	16	22	73.3	64	13	27	1014	15
4	25	16	21	74.2	74	15	21	1015	15
5	24	15	20	63.6	76	15	16	1017	15
6	23	14	18	20.1	74	13	13	1018	15
7	22	13	18	10.9	71	12	14	1018	15
8	23	13	18	17.5	70	12	16	1018	14
9	25	14	19	15.8	65	11	19	1017	15
10	26	15	20	47.8	65	13	23	1016	15
11	25	16	20	98.5	74	14	25	1015	14
12	25	16	20	76.9	72	14	26	1015	15

Calcualated average values:

```
## [1] "Annual minimum of temperatures: 13"
```

```
## [1] "Annual maximum of temperatures: 27"
```

```
## [1] "Annual mean of temperatures: 19.83"
```

```
## [1] "Annual mean precipitation: 50.52"
```

Graphs for precipitation and temperature range per month: