Ps1_Ans

Weijie Yuan

9/1/2018

Answer for Problem 3

(a)

Firstly, I create a new(temp) subdirectory within my working directory and creating my Rmd file in this new subdirectory.

```
pwd
```

```
## /Users/dogspro/Desktop/Working
```

Then, download the recent 10 years' data from https://www1.ncdc.noaa.gov/pub/data/ghcn/daily/by_year/. I use loop to download the data from 2009 to 2018.

```
for year in {2010..2018};
do
curl -O https://www1.ncdc.noaa.gov/pub/data/ghcn/daily/by_year/$year.csv.gz;
done
```

```
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                  Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0     0    0     0    0     0      0      0 --:--:--  0:00:01 --:--:--     0
  0     0    0     0    0     0      0      0 --:--:--  0:00:02 --:--:--     0
  0     0    0     0    0     0      0      0 --:--:--  0:00:03 --:--:--     0
  0     0    0     0    0     0      0      0 --:--:--  0:00:03 --:--:--     0
  0  202M    0  159k    0     0  34299      0  1:43:25  0:00:04  1:43:21 34296
  0  202M    0 1128k    0     0   193k      0  0:17:53  0:00:05  0:17:48  262k
  1  202M    1 2651k    0     0   388k      0  0:08:54  0:00:06  0:08:48  617k
  2  202M    2 4550k    0     0   587k      0  0:05:53  0:00:07  0:05:46 1080k
  3  202M    3 7190k    0     0   822k      0  0:04:12  0:00:08  0:04:04 1455k
  5  202M    5 10.6M    0     0  1112k      0  0:03:06  0:00:09  0:02:57 2137k
  7  202M    7 15.3M    0     0  1462k      0  0:02:22  0:00:10  0:02:12 2965k
 10  202M   10 20.9M    0     0  1829k      0  0:01:53  0:00:11  0:01:42 3829k
 13  202M   13 27.8M    0     0  2239k      0  0:01:32  0:00:12  0:01:20 4798k
 17  202M   17 35.4M    0     0  2640k      0  0:01:18  0:00:13  0:01:05 5818k
 21  202M   21 43.9M    0     0  3055k      0  0:01:08  0:00:14  0:00:54 6870k
 26  202M   26 53.1M    0     0  3457k      0  0:01:00  0:00:15  0:00:45 7743k
 30  202M   30 62.4M    0     0  3816k      0  0:00:54  0:00:16  0:00:38 8482k
 34  202M   34 70.8M    0     0  4088k      0  0:00:50  0:00:17  0:00:33 8800k
 39  202M   39 80.0M    0     0  4370k      0  0:00:47  0:00:18  0:00:29 9122k
 44  202M   44 89.6M    0     0  4651k      0  0:00:44  0:00:19  0:00:25 9359k
 48  202M   48 98.8M    0     0  4878k      0  0:00:42  0:00:20  0:00:22 9351k
 53  202M   53  107M    0     0  5078k      0  0:00:40  0:00:21  0:00:19 9303k
 57  202M   57  117M    0     0  5284k      0  0:00:39  0:00:22  0:00:17 9525k
 62  202M   62  126M    0     0  5470k      0  0:00:37  0:00:23  0:00:14 9599k
 66  202M   66  134M    0     0  5574k      0  0:00:37  0:00:24  0:00:13 9203k
```

```
 69  202M   69  141M    0     0  5620k      0  0:00:36  0:00:25  0:00:11 8653k
 72  202M   72  147M    0     0  5647k      0  0:00:36  0:00:26  0:00:10 8096k
 76  202M   76  154M    0     0  5698k      0  0:00:36  0:00:27  0:00:09 7584k
 79  202M   79  161M    0     0  5760k      0  0:00:36  0:00:28  0:00:08 7133k
 83  202M   83  169M    0     0  5824k      0  0:00:35  0:00:29  0:00:06 7070k
 88  202M   88  178M    0     0  5950k      0  0:00:34  0:00:30  0:00:04 7680k
 92  202M   92  188M    0     0  6069k      0  0:00:34  0:00:31  0:00:03 8352k
 97  202M   97  197M    0     0  6179k      0  0:00:33  0:00:32  0:00:01 8850k
100  202M  100  202M    0     0  6246k      0  0:00:33  0:00:33 --:--:-- 9335k
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0  192M    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0  192M    0  271k    0     0   182k      0  0:17:58  0:00:01  0:17:57  182k
  0  192M    0 1240k    0     0   504k      0  0:06:30  0:00:02  0:06:28  504k
  1  192M    1 3318k    0     0   958k      0  0:03:25  0:00:03  0:03:22  958k
  3  192M    3 6420k    0     0  1426k      0  0:02:18  0:00:04  0:02:14 1426k
  5  192M    5 10.1M    0     0  1903k      0  0:01:43  0:00:05  0:01:38 2102k
  7  192M    7 14.6M    0     0  2317k      0  0:01:25  0:00:06  0:01:19 2956k
 10  192M   10 19.8M    0     0  2726k      0  0:01:12  0:00:07  0:01:05 3819k
 13  192M   13 25.7M    0     0  3120k      0  0:01:03  0:00:08  0:00:55 4615k
 16  192M   16 30.9M    0     0  3349k      0  0:00:58  0:00:09  0:00:49 5095k
 20  192M   20 38.6M    0     0  3775k      0  0:00:52  0:00:10  0:00:42 5812k
 24  192M   24 46.5M    0     0  4157k      0  0:00:47  0:00:11  0:00:36 6532k
 28  192M   28 55.6M    0     0  4567k      0  0:00:43  0:00:12  0:00:31 7298k
 33  192M   33 65.0M    0     0  4953k      0  0:00:39  0:00:13  0:00:26 8059k
 38  192M   38 74.4M    0     0  5272k      0  0:00:37  0:00:14  0:00:23 8911k
 43  192M   43 83.7M    0     0  5539k      0  0:00:35  0:00:15  0:00:20 9239k
 48  192M   48 92.9M    0     0  5782k      0  0:00:34  0:00:16  0:00:18 9502k
 53  192M   53  102M    0     0  6011k      0  0:00:32  0:00:17  0:00:15 9636k
 57  192M   57  111M    0     0  6193k      0  0:00:31  0:00:18  0:00:13 9529k
 62  192M   62  121M    0     0  6384k      0  0:00:30  0:00:19  0:00:11 9598k
 68  192M   68  130M    0     0  6556k      0  0:00:30  0:00:20  0:00:10 9721k
 72  192M   72  140M    0     0  6697k      0  0:00:29  0:00:21  0:00:08 9713k
 77  192M   77  149M    0     0  6799k      0  0:00:29  0:00:22  0:00:07 9550k
 82  192M   82  158M    0     0  6907k      0  0:00:28  0:00:23  0:00:05 9545k
 86  192M   86  167M    0     0  7012k      0  0:00:28  0:00:24  0:00:04 9459k
 91  192M   91  176M    0     0  7089k      0  0:00:27  0:00:25  0:00:02 9272k
 95  192M   95  184M    0     0  7151k      0  0:00:27  0:00:26  0:00:01 9100k
100  192M  100  192M    0     0  7233k      0  0:00:27  0:00:27 --:--:-- 9261k
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0  191M    0  182k    0     0   148k      0  0:22:03  0:00:01  0:22:02  148k
  0  191M    0 1237k    0     0   564k      0  0:05:48  0:00:02  0:05:46  564k
  1  191M    1 3284k    0     0  1032k      0  0:03:10  0:00:03  0:03:07 1031k
  3  191M    3 6307k    0     0  1505k      0  0:02:10  0:00:04  0:02:06 1505k
  5  191M    5 10.0M    0     0  1980k      0  0:01:39  0:00:05  0:01:34 2062k
  7  191M    7 14.7M    0     0  2441k      0  0:01:20  0:00:06  0:01:14 3011k
 10  191M   10 19.7M    0     0  2813k      0  0:01:09  0:00:07  0:01:02 3799k
 13  191M   13 25.7M    0     0  3209k      0  0:01:01  0:00:08  0:00:53 4584k
```

```
 16  191M   16 32.0M    0     0  3570k      0  0:00:55  0:00:09  0:00:46 5297k
 20  191M   20 39.4M    0     0  3972k      0  0:00:49  0:00:10  0:00:39 6042k
 24  191M   24 47.7M    0     0  4374k      0  0:00:44  0:00:11  0:00:33 6765k
 28  191M   28 55.5M    0     0  4668k      0  0:00:42  0:00:12  0:00:30 7331k
 33  191M   33 64.4M    0     0  5009k      0  0:00:39  0:00:13  0:00:26 7992k
 38  191M   38 73.7M    0     0  5325k      0  0:00:36  0:00:14  0:00:22 8559k
 43  191M   43 83.0M    0     0  5603k      0  0:00:35  0:00:15  0:00:20 8925k
 48  191M   48 92.5M    0     0  5855k      0  0:00:33  0:00:16  0:00:17 9165k
 53  191M   53  102M    0     0  6082k      0  0:00:32  0:00:17  0:00:15 9533k
 58  191M   58  111M    0     0  6298k      0  0:00:31  0:00:18  0:00:13 9690k
 63  191M   63  121M    0     0  6469k      0  0:00:30  0:00:19  0:00:11 9714k
 67  191M   67  130M    0     0  6624k      0  0:00:29  0:00:20  0:00:09 9723k
 72  191M   72  140M    0     0  6773k      0  0:00:29  0:00:21  0:00:08 9745k
 77  191M   77  149M    0     0  6912k      0  0:00:28  0:00:22  0:00:06 9764k
 82  191M   82  158M    0     0  7014k      0  0:00:28  0:00:23  0:00:05 9622k
 87  191M   87  168M    0     0  7131k      0  0:00:27  0:00:24  0:00:03 9671k
 92  191M   92  178M    0     0  7237k      0  0:00:27  0:00:25  0:00:02 9709k
 97  191M   97  186M    0     0  7277k      0  0:00:27  0:00:26  0:00:01 9390k
100  191M  100  191M    0     0  7312k      0  0:00:26  0:00:26 --:--:-- 9196k
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0  192M    0  251k    0     0   185k      0  0:17:44  0:00:01  0:17:43  185k
  0  192M    0 1087k    0     0   468k      0  0:07:00  0:00:02  0:06:58  468k
  1  192M    1 2290k    0     0   697k      0  0:04:42  0:00:03  0:04:39  697k
  2  192M    2 4032k    0     0   940k      0  0:03:29  0:00:04  0:03:25  940k
  3  192M    3 6477k    0     0  1220k      0  0:02:41  0:00:05  0:02:36 1308k
  4  192M    4 9610k    0     0  1523k      0  0:02:09  0:00:06  0:02:03 1889k
  6  192M    6 13.0M    0     0  1833k      0  0:01:47  0:00:07  0:01:40 2471k
  9  192M    9 17.6M    0     0  2182k      0  0:01:30  0:00:08  0:01:22 3153k
 11  192M   11 22.8M    0     0  2516k      0  0:01:18  0:00:09  0:01:09 3871k
 14  192M   14 28.8M    0     0  2877k      0  0:01:08  0:00:10  0:00:58 4646k
 18  192M   18 36.0M    0     0  3270k      0  0:01:00  0:00:11  0:00:49 5481k
 22  192M   22 43.3M    0     0  3613k      0  0:00:54  0:00:12  0:00:42 6203k
 27  192M   27 52.0M    0     0  4016k      0  0:00:49  0:00:13  0:00:36 7075k
 31  192M   31 61.5M    0     0  4412k      0  0:00:44  0:00:14  0:00:30 7923k
 36  192M   36 69.9M    0     0  4669k      0  0:00:42  0:00:15  0:00:27 8314k
 40  192M   40 78.8M    0     0  4959k      0  0:00:39  0:00:16  0:00:23 8777k
 45  192M   45 88.3M    0     0  5207k      0  0:00:37  0:00:17  0:00:20 9054k
 46  192M   46 89.5M    0     0  4994k      0  0:00:39  0:00:18  0:00:21 7552k
 46  192M   46 90.1M    0     0  4774k      0  0:00:41  0:00:19  0:00:22 5797k
 47  192M   47 91.3M    0     0  4606k      0  0:00:42  0:00:20  0:00:22 4413k
 48  192M   48 93.3M    0     0  4488k      0  0:00:43  0:00:21  0:00:22 2955k
 49  192M   49 96.1M    0     0  4418k      0  0:00:44  0:00:22  0:00:22 1625k
 52  192M   52  100M    0     0  4409k      0  0:00:44  0:00:23  0:00:21 2224k
 54  192M   54  105M    0     0  4446k      0  0:00:44  0:00:24  0:00:20 3161k
 57  192M   57  111M    0     0  4517k      0  0:00:43  0:00:25  0:00:18 4154k
 61  192M   61  117M    0     0  4592k      0  0:00:42  0:00:26  0:00:16 5035k
 65  192M   65  125M    0     0  4701k      0  0:00:41  0:00:27  0:00:14 5967k
 68  192M   68  132M    0     0  4804k      0  0:00:41  0:00:28  0:00:13 6632k
 73  192M   73  140M    0     0  4921k      0  0:00:40  0:00:29  0:00:11 7221k
 77  192M   77  148M    0     0  5018k      0  0:00:39  0:00:30  0:00:09 7557k
```

```
 81  192M   81  157M    0     0  5160k      0  0:00:38  0:00:31  0:00:07 8152k
 86  192M   86  167M    0     0  5307k      0  0:00:37  0:00:32  0:00:05 8612k
 91  192M   91  176M    0     0  5435k      0  0:00:36  0:00:33  0:00:03 9025k
 96  192M   96  185M    0     0  5545k      0  0:00:35  0:00:34  0:00:01 9215k
100  192M  100  192M    0     0  5633k      0  0:00:35  0:00:35 --:--:-- 9553k
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0  189M    0  237k    0     0   185k      0  0:17:25  0:00:01  0:17:24  185k
  0  189M    0 1370k    0     0   610k      0  0:05:17  0:00:02  0:05:15  610k
  1  189M    1 3393k    0     0  1043k      0  0:03:05  0:00:03  0:03:02 1043k
  3  189M    3 6245k    0     0  1467k      0  0:02:12  0:00:04  0:02:08 1467k
  5  189M    5 9768k    0     0  1863k      0  0:01:44  0:00:05  0:01:39 1980k
  7  189M    7 13.7M    0     0  2249k      0  0:01:26  0:00:06  0:01:20 2782k
  9  189M    9 18.4M    0     0  2607k      0  0:01:14  0:00:07  0:01:07 3503k
 12  189M   12 24.0M    0     0  2988k      0  0:01:04  0:00:08  0:00:56 4255k
 16  189M   16 30.6M    0     0  3399k      0  0:00:57  0:00:09  0:00:48 5048k
 19  189M   19 36.9M    0     0  3693k      0  0:00:52  0:00:10  0:00:42 5613k
 23  189M   23 44.3M    0     0  4040k      0  0:00:47  0:00:11  0:00:36 6274k
 27  189M   27 52.6M    0     0  4405k      0  0:00:43  0:00:12  0:00:31 7010k
 32  189M   32 62.2M    0     0  4811k      0  0:00:40  0:00:13  0:00:27 7821k
 37  189M   37 71.4M    0     0  5141k      0  0:00:37  0:00:14  0:00:23 8360k
 42  189M   42 80.7M    0     0  5420k      0  0:00:35  0:00:15  0:00:20 8946k
 47  189M   47 90.3M    0     0  5699k      0  0:00:34  0:00:16  0:00:18 9432k
 52  189M   52 99.8M    0     0  5931k      0  0:00:32  0:00:17  0:00:15 9667k
 57  189M   57  108M    0     0  6114k      0  0:00:31  0:00:18  0:00:13 9564k
 62  189M   62  117M    0     0  6255k      0  0:00:30  0:00:19  0:00:11 9426k
 66  189M   66  126M    0     0  6409k      0  0:00:30  0:00:20  0:00:10 9430k
 72  189M   72  136M    0     0  6579k      0  0:00:29  0:00:21  0:00:08 9433k
 76  189M   76  145M    0     0  6680k      0  0:00:29  0:00:22  0:00:07 9258k
 81  189M   81  154M    0     0  6806k      0  0:00:28  0:00:23  0:00:05 9329k
 86  189M   86  163M    0     0  6915k      0  0:00:28  0:00:24  0:00:04 9455k
 91  189M   91  173M    0     0  7024k      0  0:00:27  0:00:25  0:00:02 9515k
 96  189M   96  182M    0     0  7128k      0  0:00:27  0:00:26  0:00:01 9465k
100  189M  100  189M    0     0  7174k      0  0:00:27  0:00:27 --:--:-- 9485k
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0  192M    0  149k    0     0   119k      0  0:27:25  0:00:01  0:27:24  119k
  0  192M    0 1071k    0     0   484k      0  0:06:46  0:00:02  0:06:44  483k
  1  192M    1 3047k    0     0   948k      0  0:03:27  0:00:03  0:03:24  948k
  2  192M    2 5587k    0     0  1321k      0  0:02:29  0:00:04  0:02:25 1320k
  4  192M    4 8516k    0     0  1631k      0  0:02:00  0:00:05  0:01:55 1721k
  6  192M    6 11.5M    0     0  1908k      0  0:01:43  0:00:06  0:01:37 2358k
  8  192M    8 15.5M    0     0  2185k      0  0:01:30  0:00:07  0:01:23 2930k
 10  192M   10 19.4M    0     0  2421k      0  0:01:21  0:00:08  0:01:13 3368k
 12  192M   12 24.3M    0     0  2699k      0  0:01:12  0:00:09  0:01:03 3865k
 15  192M   15 29.9M    0     0  3003k      0  0:01:05  0:00:10  0:00:55 4433k
 19  192M   19 36.5M    0     0  3334k      0  0:00:59  0:00:11  0:00:48 5106k
 22  192M   22 43.9M    0     0  3680k      0  0:00:53  0:00:12  0:00:41 5870k
```

```
 26  192M   26 51.0M    0     0  3950k      0  0:00:49  0:00:13  0:00:36 6460k
 31  192M   31 59.8M    0     0  4309k      0  0:00:45  0:00:14  0:00:31 7287k
 35  192M   35 68.7M    0     0  4626k      0  0:00:42  0:00:15  0:00:27 7950k
 40  192M   40 77.4M    0     0  4889k      0  0:00:40  0:00:16  0:00:24 8374k
 45  192M   45 86.6M    0     0  5152k      0  0:00:38  0:00:17  0:00:21 8756k
 49  192M   49 95.4M    0     0  5367k      0  0:00:36  0:00:18  0:00:18 9116k
 54  192M   54  104M    0     0  5567k      0  0:00:35  0:00:19  0:00:16 9138k
 59  192M   59  113M    0     0  5769k      0  0:00:34  0:00:20  0:00:14 9256k
 63  192M   63  122M    0     0  5933k      0  0:00:33  0:00:21  0:00:12 9325k
 68  192M   68  132M    0     0  6107k      0  0:00:32  0:00:22  0:00:10 9396k
 73  192M   73  141M    0     0  6251k      0  0:00:31  0:00:23  0:00:08 9469k
 78  192M   78  151M    0     0  6387k      0  0:00:30  0:00:24  0:00:06 9541k
 83  192M   83  160M    0     0  6522k      0  0:00:30  0:00:25  0:00:05 9559k
 88  192M   88  170M    0     0  6642k      0  0:00:29  0:00:26  0:00:03 9649k
 93  192M   93  179M    0     0  6757k      0  0:00:29  0:00:27  0:00:02 9643k
 98  192M   98  188M    0     0  6841k      0  0:00:28  0:00:28 --:--:-- 9584k
100  192M  100  192M    0     0  6878k      0  0:00:28  0:00:28 --:--:-- 9584k
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0  192M    0 14220    0     0  23149      0  2:25:06 --:--:--  2:25:06 23121
  0  192M    0  462k    0     0   292k      0  0:11:13  0:00:01  0:11:12  292k
  0  192M    0 1727k    0     0   669k      0  0:04:54  0:00:02  0:04:52  668k
  1  192M    1 3774k    0     0  1051k      0  0:03:07  0:00:03  0:03:04 1051k
  3  192M    3 6696k    0     0  1456k      0  0:02:15  0:00:04  0:02:11 1456k
  5  192M    5 10.0M    0     0  1841k      0  0:01:46  0:00:05  0:01:41 2066k
  7  192M    7 14.2M    0     0  2213k      0  0:01:28  0:00:06  0:01:22 2821k
  9  192M    9 19.0M    0     0  2566k      0  0:01:16  0:00:07  0:01:09 3540k
 12  192M   12 24.2M    0     0  2893k      0  0:01:08  0:00:08  0:01:00 4215k
 15  192M   15 29.9M    0     0  3201k      0  0:01:01  0:00:09  0:00:52 4810k
 19  192M   19 36.5M    0     0  3535k      0  0:00:55  0:00:10  0:00:45 5420k
 22  192M   22 43.2M    0     0  3828k      0  0:00:51  0:00:11  0:00:40 5954k
 26  192M   26 50.9M    0     0  4143k      0  0:00:47  0:00:12  0:00:35 6558k
 31  192M   31 59.7M    0     0  4508k      0  0:00:43  0:00:13  0:00:30 7288k
 35  192M   35 67.5M    0     0  4731k      0  0:00:41  0:00:14  0:00:27 7644k
 35  192M   35 68.7M    0     0  4508k      0  0:00:43  0:00:15  0:00:28 6563k
 36  192M   36 69.4M    0     0  4287k      0  0:00:45  0:00:16  0:00:29 5347k
 36  192M   36 70.8M    0     0  4121k      0  0:00:47  0:00:17  0:00:30 4068k
 37  192M   37 72.8M    0     0  4012k      0  0:00:49  0:00:18  0:00:31 2664k
 39  192M   39 75.5M    0     0  3950k      0  0:00:49  0:00:19  0:00:30 1656k
 41  192M   41 79.2M    0     0  3940k      0  0:00:49  0:00:20  0:00:29 2156k
 43  192M   43 83.6M    0     0  3970k      0  0:00:49  0:00:21  0:00:28 2917k
 46  192M   46 88.9M    0     0  4032k      0  0:00:48  0:00:22  0:00:26 3715k
 49  192M   49 95.0M    0     0  4126k      0  0:00:47  0:00:23  0:00:24 4551k
 52  192M   52  101M    0     0  4217k      0  0:00:46  0:00:24  0:00:22 5265k
 56  192M   56  108M    0     0  4360k      0  0:00:45  0:00:25  0:00:20 6100k
 60  192M   60  116M    0     0  4475k      0  0:00:43  0:00:26  0:00:17 6658k
 65  192M   65  125M    0     0  4647k      0  0:00:42  0:00:27  0:00:15 7429k
 69  192M   69  134M    0     0  4814k      0  0:00:40  0:00:28  0:00:12 8056k
 74  192M   74  144M    0     0  4985k      0  0:00:39  0:00:29  0:00:10 8764k
 79  192M   79  153M    0     0  5135k      0  0:00:38  0:00:30  0:00:08 9103k
 84  192M   84  162M    0     0  5274k      0  0:00:37  0:00:31  0:00:06 9515k
 89  192M   89  171M    0     0  5399k      0  0:00:36  0:00:32  0:00:04 9548k
```

```
 94  192M   94  181M    0     0  5523k      0  0:00:35  0:00:33  0:00:02 9576k
 98  192M   98  190M    0     0  5631k      0  0:00:34  0:00:34 --:--:-- 9453k
100  192M  100  192M    0     0  5656k      0  0:00:34  0:00:34 --:--:-- 9443k
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0  189M    0 49116    0     0  59870      0  0:55:14 --:--:--  0:55:14 59824
  0  189M    0  438k    0     0   244k      0  0:13:13  0:00:01  0:13:12  244k
  0  189M    0 1477k    0     0   529k      0  0:06:06  0:00:02  0:06:04  529k
  1  189M    1 2993k    0     0   793k      0  0:04:04  0:00:03  0:04:01  793k
  2  189M    2 5172k    0     0  1083k      0  0:02:58  0:00:04  0:02:54 1083k
  4  189M    4 8063k    0     0  1393k      0  0:02:19  0:00:05  0:02:14 1614k
  6  189M    6 11.3M    0     0  1714k      0  0:01:53  0:00:06  0:01:47 2243k
  8  189M    8 15.8M    0     0  2085k      0  0:01:32  0:00:07  0:01:25 2957k
 11  189M   11 21.2M    0     0  2479k      0  0:01:18  0:00:08  0:01:10 3749k
 14  189M   14 27.3M    0     0  2864k      0  0:01:07  0:00:09  0:00:58 4564k
 18  189M   18 34.0M    0     0  3239k      0  0:00:59  0:00:10  0:00:49 5375k
 22  189M   22 41.7M    0     0  3630k      0  0:00:53  0:00:11  0:00:42 6238k
 26  189M   26 50.5M    0     0  4053k      0  0:00:47  0:00:12  0:00:35 7112k
 31  189M   31 59.5M    0     0  4426k      0  0:00:43  0:00:13  0:00:30 7850k
 36  189M   36 68.6M    0     0  4758k      0  0:00:40  0:00:14  0:00:26 8460k
 41  189M   41 78.0M    0     0  5065k      0  0:00:38  0:00:15  0:00:23 9002k
 46  189M   46 87.6M    0     0  5351k      0  0:00:36  0:00:16  0:00:20 9409k
 51  189M   51 97.1M    0     0  5594k      0  0:00:34  0:00:17  0:00:17 9532k
 56  189M   56  106M    0     0  5793k      0  0:00:33  0:00:18  0:00:15 9560k
 61  189M   61  115M    0     0  5994k      0  0:00:32  0:00:19  0:00:13 9646k
 66  189M   66  125M    0     0  6162k      0  0:00:31  0:00:20  0:00:11 9628k
 70  189M   70  134M    0     0  6318k      0  0:00:30  0:00:21  0:00:09 9562k
 75  189M   75  143M    0     0  6457k      0  0:00:30  0:00:22  0:00:08 9523k
 81  189M   81  153M    0     0  6604k      0  0:00:29  0:00:23  0:00:06 9651k
 85  189M   85  162M    0     0  6726k      0  0:00:28  0:00:24  0:00:04 9623k
 90  189M   90  171M    0     0  6825k      0  0:00:28  0:00:25  0:00:03 9576k
 95  189M   95  180M    0     0  6921k      0  0:00:27  0:00:26  0:00:01 9546k
100  189M  100  189M    0     0  7006k      0  0:00:27  0:00:27 --:--:-- 9568k
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0  111M    0  188k    0     0   161k      0  0:11:43  0:00:01  0:11:42  161k
  1  111M    1 1165k    0     0   543k      0  0:03:29  0:00:02  0:03:27  543k
  2  111M    2 2633k    0     0   844k      0  0:02:14  0:00:03  0:02:11  844k
  4  111M    4 4696k    0     0  1144k      0  0:01:39  0:00:04  0:01:35 1144k
  6  111M    6 7532k    0     0  1468k      0  0:01:17  0:00:05  0:01:12 1498k
  9  111M    9 10.6M    0     0  1776k      0  0:01:04  0:00:06  0:00:58 2156k
 12  111M   12 14.2M    0     0  2052k      0  0:00:55  0:00:07  0:00:48 2704k
 16  111M   16 18.6M    0     0  2350k      0  0:00:48  0:00:08  0:00:40 3292k
 20  111M   20 23.0M    0     0  2595k      0  0:00:43  0:00:09  0:00:34 3787k
 25  111M   25 28.1M    0     0  2844k      0  0:00:40  0:00:10  0:00:30 4261k
 30  111M   30 33.9M    0     0  3130k      0  0:00:36  0:00:11  0:00:25 4792k
 35  111M   35 39.6M    0     0  3335k      0  0:00:34  0:00:12  0:00:22 5137k
 41  111M   41 45.9M    0     0  3587k      0  0:00:31  0:00:13  0:00:18 5591k
 47  111M   47 53.3M    0     0  3870k      0  0:00:29  0:00:14  0:00:15 6189k
```

```
 55  111M   55 61.8M    0     0  4194k      0  0:00:27  0:00:15  0:00:12 6931k
 63  111M   63 71.0M    0     0  4519k      0  0:00:25  0:00:16  0:00:09 7603k
 72  111M   72 80.7M    0     0  4835k      0  0:00:23  0:00:17  0:00:06 8525k
 80  111M   80 89.7M    0     0  5077k      0  0:00:22  0:00:18  0:00:04 8989k
 89  111M   89 99.0M    0     0  5307k      0  0:00:21  0:00:19  0:00:02 9363k
 97  111M   97  108M    0     0  5508k      0  0:00:20  0:00:20 --:--:-- 9478k
100  111M  100  111M    0     0  5572k      0  0:00:20  0:00:20 --:--:-- 9501k
```

To report the number of observations for Death Valley in each year by printing the information to the screen, I search the station ID codes of Death Valley is USC00042319.

```
# unzip .gz files
for year in {2010..2018}; do gunzip $year.csv.gz; done
```

```
# report the number of observation each year
for year in {2010..2018}; do echo "${year} $(wc -l $year.csv)"; done
```

```
## 2010  37840433 2010.csv
## 2011  34827969 2011.csv
## 2012  35134778 2012.csv
## 2013  35314935 2013.csv
## 2014  34599683 2014.csv
## 2015  35233244 2015.csv
## 2016  35384539 2016.csv
## 2017  34748555 2017.csv
## 2018  20513640 2018.csv
```

(b)

Firstly, I download the ghcnd-stations.txt including information about station ID codes and check the data structure by looking over the first 5 rows.

```
curl -O https://www1.ncdc.noaa.gov/pub/data/ghcn/daily/ghcnd-stations.txt
head -n 5 ghcnd-stations.txt
```

```
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                  Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0 8959k    0 14231    0     0  20408      0  0:07:29 --:--:--  0:07:29 20388
  3 8959k    3  297k    0     0   178k      0  0:00:50  0:00:01  0:00:49  178k
 12 8959k   12 1110k    0     0   417k      0  0:00:21  0:00:02  0:00:19  417k
 24 8959k   24 2157k    0     0   578k      0  0:00:15  0:00:03  0:00:12  578k
 38 8959k   38 3415k    0     0   720k      0  0:00:12  0:00:04  0:00:08  720k
 54 8959k   54 4876k    0     0   853k      0  0:00:10  0:00:05  0:00:05  969k
 74 8959k   74 6633k    0     0   991k      0  0:00:09  0:00:06  0:00:03 1261k
 98 8959k   98 8860k    0     0  1156k      0  0:00:07  0:00:07 --:--:-- 1548k
100 8959k  100 8959k    0     0  1163k      0  0:00:07  0:00:07 --:--:-- 1711k
## ACW00011604  17.1167  -61.7833   10.1    ST JOHNS COOLIDGE FLD
## ACW00011647  17.1333  -61.7833   19.2    ST JOHNS
## AE000041196  25.3330   55.5170   34.0    SHARJAH INTER. AIRP            GSN     41196
## AEM00041194  25.2550   55.3640   10.4    DUBAI INTL                             41194
## AEM00041217  24.4330   54.6510   26.8    ABU DHABI INTL                         41217
```

Then I extract the station ID code of Death Valley for the ghcn-stations.txt by match the "DEATH VALLEY"

and cut the first part. Assign the station ID to a variable named id_of_deathvalley.

```
id_of_deathvalley=$(grep "DEATH VALLEY" ghcnd-stations.txt | cut -d' ' -f1)
echo $id_of_deathvalley
#After that, I use mkdir to bulid a new subdirectory and extract all lines matching
#station ID code, TMAX and March from original .csv files. And put these matched items
#into new corresponding .csv files. Finally I put all these generated .csv files
#together.
mkdir subset
for year in {2010..2018}; do
grep "$id_of_deathvalley" $year.csv |
grep "TMAX" |
grep "${year}03" > subset/subset_$year.csv; done
cd subset
cat *.csv > sum_subset.csv
```

```
## USC00042319
```

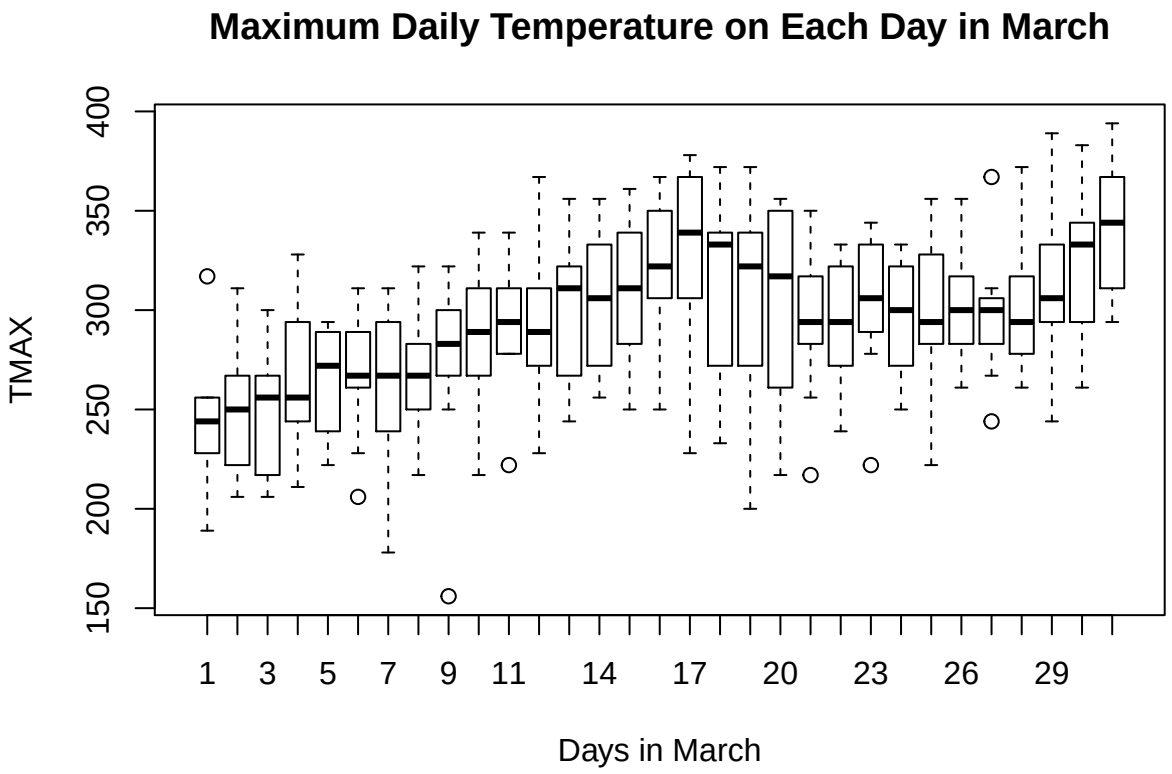
```
# Move sum_subset.csv out to current working directory and check its data structure by
#looking over its first ten items.
mv subset/sum_subset.csv .
head -n 10 sum_subset.csv
```

```
## USC00042319,20100301,TMAX,244,,,0,0800
## USC00042319,20100302,TMAX,256,,,0,0800
## USC00042319,20100303,TMAX,267,,,0,0800
## USC00042319,20100304,TMAX,256,,,0,0800
## USC00042319,20100305,TMAX,233,,,0,0800
## USC00042319,20100306,TMAX,228,,,0,0800
## USC00042319,20100307,TMAX,233,,,0,0800
## USC00042319,20100308,TMAX,250,,,0,0800
## USC00042319,20100309,TMAX,250,,,0,0800
## USC00042319,20100310,TMAX,217,,,0,0800
```

(c)

Extract all the TMAX each day in March from the fourth column of sum_subset. Convert the data into a matrix of which shape is 9*31, which means 31 days in March and 9 years data respectively.

```
data <- read.csv("sum_subset.csv",header = FALSE)
TMAX_index = rep(1:31,times=9)
boxplot(data[,4] ~TMAX_index,main="Maximum Daily Temperature on Each Day in March",
        xlab ="Days in March",ylab="TMAX")
```



(d)

I build a bash function to match user's interst. There are four input parameters and one output information. Fours inputs include years, month, station and weather. "Years" parameter has an array-like structure and the most advantage of using array is that you can select years that are not successive. For example, a user just want to compare the weather feature between 1900 and 2018. This structure can be used to save a lot time which is supposed to download the data from 1901-2017. The "month" parameter can be numeric or string with or without quote. The "weather" function can also be numeric or string. And spcial case is "Station" parameter, because users may input station name with spacing between words, which is unavoidable. So this parameter should be enter with quotes on both sides of station name.

```
function get_weather(){
if [ "$1" == "-h" ]; then
  echo "This function is design to output the maximum daily temperature of years and
  month according to user's interest.
  Input:
  years: years of interest in format '({year1,year2,...})[@]'
  month: numeric value, from 01 to 12
  station: a string of the name of the weather station (UPPER CASE)
  weather: the weather features of interest.
    the input of weather can only be below:
    PRCP = Precipitation(tenths of mm)
    SNOW = Snowfall (mm)
    SNWD = Snow depth (mm)
    TMAX = Maximum temperature (tenths of degree in C)
    TMIN = Minimum temperature (tenths of degree in C)
    TAVG = Average temperature
  Output:
```

```
  information containing user's interest"
  exit 0
fi
# check if the number of argument is legal. If illegal, return the error message and exit
# the fuction.
if [ "$#" -ne 4 ]; then
echo "Error: Invalid arguments please pass exactly four arguments"
exit 1
fi
# check if the weather features is legal. If illegal, return the error message and exit
# the function.
if ! [ "$4" == "PRCP" ] && ! [ "$4" == "SNOW" ] &&
! [ "$4" == "SNWD" ] && ! [ "$4" == "TMAX" ] &&
! [ "$4" == "TMIN" ] && ! [ "$4" == "TAVG" ]; then
echo "Error: Wrong weather feature of interest"
exit 1
fi
#creat new temporary directory
mkdir temp;
cd temp;
arr=("${!1}");
# loop by years provided in the first argument.
# download and unzip
for year in "${arr[@]}"
do
curl -O https://www1.ncdc.noaa.gov/pub/data/ghcn/daily/by_year/$year.csv.gz;
gzip -d $year.csv.gz;
done
# download the .txt file including the information about station names and ID codes.
curl -O https://www1.ncdc.noaa.gov/pub/data/ghcn/daily/ghcnd-stations.txt
# match the input name with corresponding id code
id=$(grep "$3" ghcnd-stations.txt | cut -d' ' -f1)
# check if the the input name can identify a single weather station
# if not, return error message and exit the function
if ! [ -n $id ]
then
echo "Error: Can not identify a single weather station"
cd ..
rm -rf temp
exit 1
fi
# put the matched information together and return the result
mkdir subset;
for year in "${arr[@]}";
do
grep "$id" $year.csv |
grep "$4" |
grep "${year}$2" > subset/subset_$year.csv;
done
cd subset
cat *.csv > sum_subset.csv
cat sum_subset.csv
# remove the raw downloaded data files
```

```
cd ..
cd ..
rm -rf temp
}

#get_weather -h
a=({2010,2011}) #users can add any years of interest
get_weather a[@] 04 "DEATH VALLEY" TMIN
```

```
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                  Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0  202M    0  346k    0     0   257k      0  0:13:26  0:00:01  0:13:25  257k
  0  202M    0 1557k    0     0   651k      0  0:05:19  0:00:02  0:05:17  651k
  1  202M    1 3542k    0     0  1054k      0  0:03:17  0:00:03  0:03:14 1053k
  3  202M    3 6596k    0     0  1518k      0  0:02:16  0:00:04  0:02:12 1517k
  4  202M    4  9.9M    0     0  1908k      0  0:01:48  0:00:05  0:01:43 2043k
  7  202M    7 14.3M    0     0  2324k      0  0:01:29  0:00:06  0:01:23 2880k
  9  202M    9 19.3M    0     0  2692k      0  0:01:17  0:00:07  0:01:10 3678k
 12  202M   12 24.9M    0     0  3065k      0  0:01:07  0:00:08  0:00:59 4422k
 15  202M   15 31.5M    0     0  3449k      0  0:01:00  0:00:09  0:00:51 5126k
 18  202M   18 38.5M    0     0  3812k      0  0:00:54  0:00:10  0:00:44 5850k
 23  202M   23 46.8M    0     0  4233k      0  0:00:49  0:00:11  0:00:38 6654k
 27  202M   27 55.5M    0     0  4606k      0  0:00:45  0:00:12  0:00:33 7419k
 31  202M   31 64.6M    0     0  4960k      0  0:00:41  0:00:13  0:00:28 8122k
 36  202M   36 73.6M    0     0  5257k      0  0:00:39  0:00:14  0:00:25 8643k
 40  202M   40 83.0M    0     0  5544k      0  0:00:37  0:00:15  0:00:22 9131k
 45  202M   45 91.6M    0     0  5742k      0  0:00:36  0:00:16  0:00:20 9164k
 49  202M   49  101M    0     0  5975k      0  0:00:34  0:00:17  0:00:17 9354k
 54  202M   54  109M    0     0  6124k      0  0:00:33  0:00:18  0:00:15 9228k
 58  202M   58  119M    0     0  6310k      0  0:00:32  0:00:19  0:00:13 9333k
 63  202M   63  128M    0     0  6457k      0  0:00:32  0:00:20  0:00:12 9259k
 67  202M   67  137M    0     0  6591k      0  0:00:31  0:00:21  0:00:10 9358k
 72  202M   72  146M    0     0  6723k      0  0:00:30  0:00:22  0:00:08 9319k
 76  202M   76  156M    0     0  6848k      0  0:00:30  0:00:23  0:00:07 9504k
 81  202M   81  165M    0     0  6956k      0  0:00:29  0:00:24  0:00:05 9451k
 85  202M   85  174M    0     0  7039k      0  0:00:29  0:00:25  0:00:04 9409k
 90  202M   90  182M    0     0  7112k      0  0:00:29  0:00:26  0:00:03 9342k
 94  202M   94  191M    0     0  7173k      0  0:00:28  0:00:27  0:00:01 9181k
 98  202M   98  200M    0     0  7227k      0  0:00:28  0:00:28 --:--:-- 8996k
100  202M  100  202M    0     0  7257k      0  0:00:28  0:00:28 --:--:-- 8962k
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                  Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0  192M    0 67252    0     0  69031      0  0:48:45 --:--:--  0:48:45 68976
  0  192M    0  635k    0     0   326k      0  0:10:04  0:00:01  0:10:03  326k
  0  192M    0 1901k    0     0   649k      0  0:05:03  0:00:02  0:05:01  649k
  1  192M    1 3667k    0     0   933k      0  0:03:31  0:00:03  0:03:28  933k
  3  192M    3 6284k    0     0  1272k      0  0:02:34  0:00:04  0:02:30 1272k
  4  192M    4 9753k    0     0  1639k      0  0:02:00  0:00:05  0:01:55 1946k
  7  192M    7 13.5M    0     0  2001k      0  0:01:38  0:00:06  0:01:32 2656k
```

```
  9  192M    9 18.2M    0     0  2351k      0  0:01:23  0:00:07  0:01:16 3348k
 12  192M   12 24.1M    0     0  2774k      0  0:01:11  0:00:08  0:01:03 4219k
 16  192M   16 30.8M    0     0  3179k      0  0:01:02  0:00:09  0:00:53 5065k
 19  192M   19 38.4M    0     0  3602k      0  0:00:54  0:00:10  0:00:44 5948k
 24  192M   24 46.6M    0     0  4002k      0  0:00:49  0:00:11  0:00:38 6777k
 28  192M   28 55.4M    0     0  4388k      0  0:00:44  0:00:12  0:00:32 7618k
 33  192M   33 64.2M    0     0  4722k      0  0:00:41  0:00:13  0:00:28 8204k
 37  192M   37 73.0M    0     0  5009k      0  0:00:39  0:00:14  0:00:25 8643k
 42  192M   42 82.4M    0     0  5303k      0  0:00:37  0:00:15  0:00:22 9023k
 47  192M   47 92.1M    0     0  5576k      0  0:00:35  0:00:16  0:00:19 9332k
 52  192M   52  101M    0     0  5777k      0  0:00:34  0:00:17  0:00:17 9367k
 57  192M   57  110M    0     0  5990k      0  0:00:32  0:00:18  0:00:14 9521k
 62  192M   62  120M    0     0  6168k      0  0:00:31  0:00:19  0:00:12 9626k
 67  192M   67  129M    0     0  6338k      0  0:00:31  0:00:20  0:00:11 9628k
 71  192M   71  138M    0     0  6463k      0  0:00:30  0:00:21  0:00:09 9458k
 76  192M   76  147M    0     0  6609k      0  0:00:29  0:00:22  0:00:07 9595k
 81  192M   81  157M    0     0  6734k      0  0:00:29  0:00:23  0:00:06 9550k
 86  192M   86  166M    0     0  6842k      0  0:00:28  0:00:24  0:00:04 9532k
 91  192M   91  175M    0     0  6941k      0  0:00:28  0:00:25  0:00:03 9471k
 96  192M   96  184M    0     0  7032k      0  0:00:28  0:00:26  0:00:02 9534k
100  192M  100  192M    0     0  7106k      0  0:00:27  0:00:27 --:--:-- 9471k
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0 8959k    0 33127    0     0  45035      0  0:03:23 --:--:--  0:03:23 45009
  6 8959k    6  602k    0     0   353k      0  0:00:25  0:00:01  0:00:24  353k
 22 8959k   22 2008k    0     0   743k      0  0:00:12  0:00:02  0:00:10  743k
 45 8959k   45 4079k    0     0  1097k      0  0:00:08  0:00:03  0:00:05 1097k
 72 8959k   72 6501k    0     0  1387k      0  0:00:06  0:00:04  0:00:02 1387k
100 8959k  100 8959k    0     0  1632k      0  0:00:05  0:00:05 --:--:-- 1879k
## USC00042319,20100401,TMIN,78,,,0,0800
## USC00042319,20100402,TMIN,100,,,0,0800
## USC00042319,20100403,TMIN,139,,,0,0800
## USC00042319,20100404,TMIN,128,,,0,0800
## USC00042319,20100405,TMIN,178,,,0,0800
## USC00042319,20100406,TMIN,111,,,0,0800
## USC00042319,20100407,TMIN,161,,,0,0800
## USC00042319,20100408,TMIN,133,,,0,0800
## USC00042319,20100409,TMIN,144,,,0,0800
## USC00042319,20100410,TMIN,183,,,0,0800
## USC00042319,20100411,TMIN,200,,,0,0800
## USC00042319,20100412,TMIN,156,,,0,0800
## USC00042319,20100413,TMIN,111,,,0,0800
## USC00042319,20100414,TMIN,128,,,0,0800
## USC00042319,20100415,TMIN,139,,,0,0800
## USC00042319,20100416,TMIN,167,,,0,0800
## USC00042319,20100417,TMIN,194,,,0,0800
## USC00042319,20100418,TMIN,183,,,0,0800
## USC00042319,20100419,TMIN,211,,,0,0800
## USC00042319,20100420,TMIN,183,,,0,0800
## USC00042319,20100421,TMIN,78,,,0,0800
## USC00042319,20100422,TMIN,100,,,0,0800
## USC00042319,20100423,TMIN,100,,,0,0800
```

```
## USC00042319,20100424,TMIN,167,,,0,0800
## USC00042319,20100425,TMIN,194,,,0,0800
## USC00042319,20100426,TMIN,211,,,0,0800
## USC00042319,20100427,TMIN,211,,,0,0800
## USC00042319,20100428,TMIN,217,,,0,0800
## USC00042319,20100429,TMIN,156,,,0,0800
## USC00042319,20100430,TMIN,128,,,0,0800
## USC00042319,20110401,TMIN,183,,,7,0800
## USC00042319,20110402,TMIN,206,,,7,0800
## USC00042319,20110403,TMIN,222,,,7,0800
## USC00042319,20110404,TMIN,167,,,7,0800
## USC00042319,20110405,TMIN,117,,,7,0800
## USC00042319,20110406,TMIN,106,,,7,0800
## USC00042319,20110407,TMIN,183,,,7,0800
## USC00042319,20110408,TMIN,106,,,7,0800
## USC00042319,20110409,TMIN,94,,,7,0800
## USC00042319,20110410,TMIN,100,,,7,0800
## USC00042319,20110411,TMIN,117,,,7,0800
## USC00042319,20110412,TMIN,128,,,7,0800
## USC00042319,20110413,TMIN,178,,,7,0800
## USC00042319,20110414,TMIN,161,,,7,0800
## USC00042319,20110415,TMIN,133,,,7,0800
## USC00042319,20110416,TMIN,144,,,7,0800
## USC00042319,20110417,TMIN,194,,,7,0800
## USC00042319,20110418,TMIN,228,,,7,0800
## USC00042319,20110419,TMIN,228,,,7,0800
## USC00042319,20110420,TMIN,222,,,7,0800
## USC00042319,20110421,TMIN,217,,,7,0800
## USC00042319,20110422,TMIN,194,,,7,0800
## USC00042319,20110423,TMIN,217,,,7,0800
## USC00042319,20110424,TMIN,217,,,7,0800
## USC00042319,20110425,TMIN,183,,,7,0800
## USC00042319,20110426,TMIN,200,,,7,0800
## USC00042319,20110427,TMIN,178,,,7,0800
## USC00042319,20110428,TMIN,172,,,7,0800
## USC00042319,20110429,TMIN,206,,,7,0800
## USC00042319,20110430,TMIN,139,,,7,0800
```

Answer for Problem 4

Use grep and sed function extract all the file name with .txt from the HTML index file. Download all of them and return status message including the name of file to the user.

```
for file in $(curl -s https://www1.ncdc.noaa.gov/pub/data/ghcn/daily/ |
                  grep href |
                  sed 's/.*href="//' |
                  sed 's/".*//' |
                  grep '^[a-zA-Z].*'| grep '.txt'); do
    echo -e "\n $file"
    curl -O https://www1.ncdc.noaa.gov/pub/data/ghcn/daily/$file
done
```

```
##
##  ghcnd-countries.txt
```

```
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
100  3670  100  3670    0     0   8044      0 --:--:-- --:--:-- --:--:--  8048
##
##  ghcnd-inventory.txt
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0 26.6M    0  128k    0     0   122k      0  0:03:43  0:00:01  0:03:42  122k
  3 26.6M    3  948k    0     0   451k      0  0:01:00  0:00:02  0:00:58  451k
 10 26.6M   10 2737k    0     0   887k      0  0:00:30  0:00:03  0:00:27  887k
 21 26.6M   21 5745k    0     0  1417k      0  0:00:19  0:00:04  0:00:15 1417k
 39 26.6M   39 10.4M    0     0  2118k      0  0:00:12  0:00:05  0:00:07 2163k
 63 26.6M   63 17.0M    0     0  2888k      0  0:00:09  0:00:06  0:00:03 3468k
 91 26.6M   91 24.4M    0     0  3546k      0  0:00:07  0:00:07 --:--:-- 4853k
100 26.6M  100 26.6M    0     0  3724k      0  0:00:07  0:00:07 --:--:-- 5785k
##
##  ghcnd-states.txt
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
100  1086  100  1086    0     0   2301      0 --:--:-- --:--:-- --:--:--  2305
##
##  ghcnd-stations.txt
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  2 8959k    2  237k    0     0   192k      0  0:00:46  0:00:01  0:00:45  192k
 14 8959k   14 1315k    0     0   592k      0  0:00:15  0:00:02  0:00:13  592k
 38 8959k   38 3425k    0     0  1063k      0  0:00:08  0:00:03  0:00:05 1063k
 71 8959k   71 6370k    0     0  1500k      0  0:00:05  0:00:04  0:00:01 1500k
100 8959k  100 8959k    0     0  1802k      0  0:00:04  0:00:04 --:--:-- 1904k
##
##  ghcnd-version.txt
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
100   270  100   270    0     0    616      0 --:--:-- --:--:-- --:--:--   616
##
##  mingle-list.txt
##   % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                 Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
  2 3707k    2 83263    0     0  97518      0  0:00:38 --:--:--  0:00:38 97497
```

```
 21 3707k   21  807k    0     0   446k      0  0:00:08  0:00:01  0:00:07  446k
 73 3707k   73 2737k    0     0   976k      0  0:00:03  0:00:02  0:00:01  975k
100 3707k  100 3707k    0     0  1180k      0  0:00:03  0:00:03 --:--:-- 1180k
##
##  readme.txt
##    % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                   Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
100 26498  100 26498    0     0  43188      0 --:--:-- --:--:-- --:--:-- 43226
##
##  status.txt
##    % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
##                                   Dload  Upload   Total   Spent    Left  Speed
##
  0     0    0     0    0     0      0      0 --:--:-- --:--:-- --:--:--     0
100 31860  100 31860    0     0  48821      0 --:--:-- --:--:-- --:--:-- 48865
```

Answer for Problem 5

(b)

Download the "reticulate" package and load it into R workplace.

```
library(reticulate)
use_python('/Users/dogspro/anaconda3/bin/python')
```

Read a dataset into R and check its data structure.

```
data <- read.csv("/Users/dogspro/Desktop/Working/stat243-S01/forestfires.csv",header = TRUE)
data[1:10,]
```

```
##    X Y month day FFMC   DMC    DC  ISI temp RH wind rain area
## 1  7 5   mar fri 86.2  26.2  94.3  5.1  8.2 51  6.7  0.0    0
## 2  7 4   oct tue 90.6  35.4 669.1  6.7 18.0 33  0.9  0.0    0
## 3  7 4   oct sat 90.6  43.7 686.9  6.7 14.6 33  1.3  0.0    0
## 4  8 6   mar fri 91.7  33.3  77.5  9.0  8.3 97  4.0  0.2    0
## 5  8 6   mar sun 89.3  51.3 102.2  9.6 11.4 99  1.8  0.0    0
## 6  8 6   aug sun 92.3  85.3 488.0 14.7 22.2 29  5.4  0.0    0
## 7  8 6   aug mon 92.3  88.9 495.6  8.5 24.1 27  3.1  0.0    0
## 8  8 6   aug mon 91.5 145.4 608.2 10.7  8.0 86  2.2  0.0    0
## 9  8 6   sep tue 91.0 129.5 692.6  7.0 13.1 63  5.4  0.0    0
## 10 7 5   sep sat 92.5  88.0 698.6  7.1 22.8 40  4.0  0.0    0
```

Slightly process the data from R in python using "pandas". One can read data named 'dataset' from R to python using "r.dataset" and read data from python to R using "py$dataset". Furthermore, one can change environment from R to python in R console using "repl_python()" and back to R using "exit".

```
import pandas
fires = r.data
fires_filter = fires[fires['month']=='sep'][['temp','DMC']]
```

Using ggplot to output visual result to users.

```
library(ggplot2)
ggplot(py$fires,aes(DMC,temp)) +geom_point()
```