ME 397M: Applied Engineering Data Analysis, Optimization and Visualization, Homework #1

Instructor: Joshua Rhodes, PhD

Due: Sunday, September 30, 2018 at 11:59pm

The deliverables for the two tasks below will be two R functions, one for each task. To turn in your assignment, 1) put both R functions in the same folder and name that folder: LAST_NAME_HWK1 2) compress that folder to a zip file (LAST_NAME_HWK1.zip), and 3) upload that zipped folder to canvas.

1. Create a function in R that downloads and formats data from the PVWATTS website given an input latitude, longitude, capacity, and array type. The output of the function should be a dataframe that has one column with a fractional value of the solar output divided by the capacity in watts. **You will need to sign up for your own API key to replace the 'XXXXXXXXX…' in the URL below -- get one on the documentation web page**: https://developer.nrel.gov/docs/solar/pvwatts/v5/

   For example, the following URL accesses data for a PV solar unit with a capacity of 100 kW located in Oklahoma City that has ``Fixed - Roof Mounted" as its array type.

https://developer.nrel.gov/api/pvwatts/v5.json?api_key=XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX&lat=35.5&lon=-97.5&system_capacity=0.1&array_type=1&module_type=1&losses=10&azimuth=180&tilt=35.5&timeframe=hourly

   The function will look something like this:

```
LAST_NAME_PVWATTS_HWK1 <- function(lat, long, capacity, array_type) {
... create URL from inputs ...
... download data ...
... format data to express solar output as a fraction ...
... return data to function call ...
}
```

   The deliverable will be a function that I will be able to run on my own console and get to provide results. For instance, a correctly functioning function would return the following. Note that the output to use from PVWATTS should be hourly data of AC generation.

```
> source('WHITE_PVWATTS_HW1.R')
> data <- WHITE_PVWATTS_HW1(lat = 35.5, long = -97.5, capacity = 10, array_type = 1)
> summary(data)
 Solar_Fraction
 Min.    :0.000000
 1st Qu.:0.000000
 Median :0.008245
 Mean    :0.186274
 3rd Qu.:0.335682
 Max.    :0.909090
```

2. The R script called `RHODES_TXPUC.R' on Canvas scrapes all of the PDFs linked from:

   https://www.puc.texas.gov/industry/electric/rates/RESrate/RESratearc.aspx

   and creates a dataframe of all the results.

   For this assignment you will make a similar function that generates the output from a single month. Your function will take inputs of month and year, read in the correct PDF from the Texas Public Utilities Commission website and then write a CSV file of the results.

   The function will look something like this:

```
LAST_NAME_TXPUC_SINGLE_HWK1 <- function(month, year) {
... create URL from inputs ...
... download linked PDF & clean data ...
... format data from tables into a dataframe ...
... write CSV file of results ...
}
```

   Below is a screenshot of what the input and output of the function should look like for the given inputs, but should be able to take other valid inputs and still output results.

```
> source('RHODES_TXPUC_SINGLE_HWK1.R')
> RHODES_TXPUC_SINGLE_HWK1(month = 'Jun', year = '07')
trying URL 'https://www.puc.texas.gov/industry/electric/rates/RESrate/rate07/Jun07Rates.pdf'
Content type 'application/pdf' length 170630 bytes (166 KB)
==================================================
downloaded 166 KB

> month_rates <- read.csv('RHODES_TXPUC_HWK1_Jun_07.csv')
> summary(month_rates)
             Plan        kWh500          kWh1000         kWh1500         kWh2000             year     month
 100% Wind     :  7   Min.   :12.20   Min.   :11.90   Min.   :11.90   Min.   :11.90   Min.   :7   Jun:149
 100% Renewable:  5   1st Qu.:13.70   1st Qu.:13.23   1st Qu.:13.14   1st Qu.:13.06   1st Qu.:7
 Deep Discount :  5   Median :14.29   Median :13.80   Median :13.70   Median :13.62   Median :7
 DYNOWATT      :  5   Mean   :14.49   Mean   :14.01   Mean   :13.87   Mean   :13.80   Mean   :7
 GEXA ENERGY   :  5   3rd Qu.:15.04   3rd Qu.:14.54   3rd Qu.:14.30   3rd Qu.:14.20   3rd Qu.:7
 Green Power   :  5   Max.   :18.57   Max.   :17.91   Max.   :17.70   Max.   :17.59   Max.   :7
 (Other)       :117
```

   The function should be able take different months and years and still provide output.