ME 397M: Applied Engineering Data Analysis, Optimization and Visualization, Midterm

Instructor: Joshua Rhodes, PhD

Due: Sunday, November 4, 2018 11:59pm CST

Task 1: Create two tables on our database at TACC, one of the 2016 EIA data, and the other from the EIA 2016 Form 860M data. Then create a PostgreSQL function that merges the two tables.

1.1 Table 1 should be named "LASTNAME_FIRSTNAME_EIA_PP" and should contain the data found in the CSV file "RHODES_JOSHUA_EIA_2016_data.csv"

1.2 Table 2 should be named "LASTNAME_FIRSTNAME_EIA_LOC" and should contain the data found in the CSV file "EIA_F860M_LL.csv"

1.3 The PostgreSQL script you generate should merge the tables in Plant ID/Plant Code and return data that includes Plant ID, Technology, Nameplate capacity, latitude, and longitude.

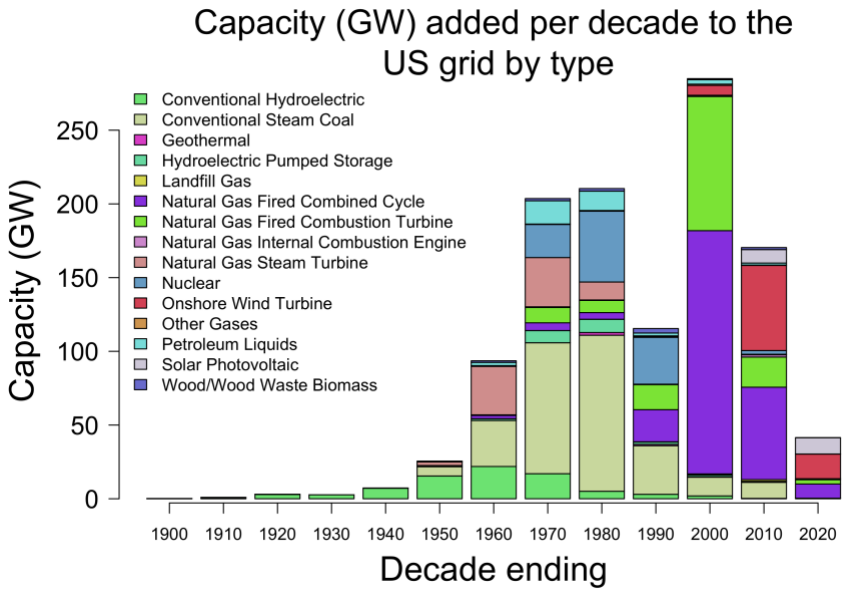
Task 1 deliverables: two tables Task 1.1 and Task 1.2 on the class database and a query file (.sql) that includes the script used to merge and generate the data explained in Task 1.3

You will need to give Philip and I access to the databases you create by running the code twice, once for each table you create:

GRANT SELECT ON public.*table* TO joshdr;

GRANT SELECT ON public.*table* TO whiphi92;

Task 2: Create and run a job on the TACC resource Maverick. You will create all the necessary files to run a R function on Maverick (hint: 3). The job you submit will read in the data from the CSV file "RHODES_JOSHUA_EIA_2016_data.csv" and will generate a PDF of a bar chart of total power plant capacity added by decade that looks something like this:



Note: the colors might vary in your final figure, and that is ok. Hint: see lecture documents on data visualization and using TACC.

Deliverables for Task 2: Three (3) files needed to submit a job on TACC that I can run myself. Hint: *do not hardcode file paths in your files* – assume that the R function that creates the figure will be in the same directory as the input CSV data file.

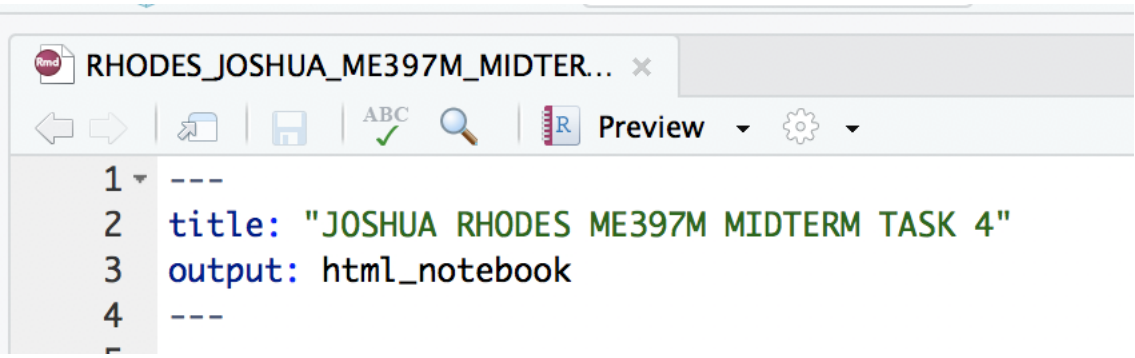
Task 3: In groups of 3, you will create a simple GitHub repository where you will each edit a document a few times. For instance, one person will create a repository and a README.md file with some text. The second person will pull that repository to their own local machine, edit the README file by adding some lines to it, and deleting others. The third person will do the same as the second and then each will do it one more time. Thus each person in the group will edit the document at least twice. You will need to invite Philip (whiphi92) and I (joshdr83) to the repository so that we can view the changes made to the document from each user.

Deliverables for Task 3: Evidence if you all having done Task 3 by us observing the history of the README file on the Repository. If you do not know how to use Github, you can take this quick tutorial that can help get you going.

Task 4: Using GIS data and tools, you will figure out the technical potential for producing offshore wind energy in the Texas Gulf Coast. The area we are interested in are GLO State Submerged Lands, which are state-owned tracts in offshore waters and coastal bays that are owned and leased by the Texas General Land Office or GLO. Using these data and NREL's Gulf of Mexico 90-m Windspeed Offshore Wind, you will calculate how much energy might be produced per year on the GLO lands (a rough conversion of the data in this chart defines a relationship between average wind speed and average capacity factor: cap_fac = wind_speed*5.26 - 8.38). However, you will also need to restrict building of wind turbines so

that they are not within 3 miles of shoreline as defined in the US_COUNTY_SHPFILE data. Hint: the US_COUNTY_SHPFILE data are counties, but you will just need a map of the entire state of Texas in this analysis. We will assume that the density of wind turbines will be that of the new Walney Extension wind farm in England (413 acres/turbine) and that the capacity of the turbines are 8.25 MW each.

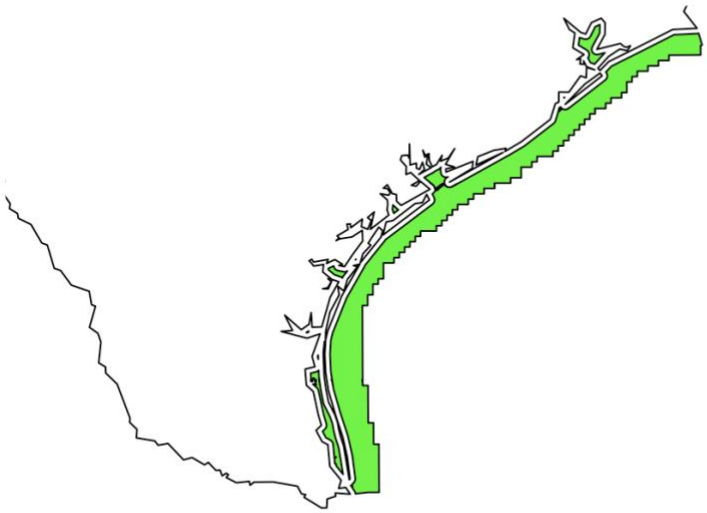
Deliverables for Task 4: An R Notebook with the title: "FIRST_NAME LAST_NAME ME379M MIDTERM TASK 4" and named: "RHODES_JOSHUA_ME397M_MIDTERM_TASK_4.Rmd" that shows the progression of your analysis. All lines of code should be commented.



You should have the notebook display the following:

1) A map of the state of Texas (not counties) that shows the leftover GLO tracts (after the 3mi buffer is removed) where the plotting extent is limited to the max x and y values of the GLO tracts (see GIS lecture 2 code notes), it should look something like this:

Texas Gulf Coast with GLO State Submerged Lands considered for wind development



2) The first 6 rows of the final dataframe that has all the GLO tract information as well as the calculated energy:

```
head(wind.tracts.full)

print(paste('The TX Gulf Coast GLO State Submerged Lands could produce up to: ', prettyNum(round(total_energy, 0),big.mark=",",scientific=FALSE), ' TWh per year.', sep = ''))
```

data.frame 6 x 16 | R Console

WATERBODY <chr>	SUB_SDE_NU <chr>	Shape_STAr <dbl>	Shape_STLe <dbl>	ID <chr>	GRIDCODE <chr>	Speed_90 <dbl>	cap_fac <dbl>	num_turbines <dbl>	annual_energy_produced <dbl>
Gulf of Mexico	990021543	3865009	8333.645	1657	7625	7.625	0.317275	1.743818	39984.82
Gulf of Mexico	990021544	3864914	8333.863	1657	7625	7.625	0.317275	1.743850	39985.54
Gulf of Mexico	990021545	3865310	8334.110	1657	7625	7.625	0.317275	1.743584	39979.44
Gulf of Mexico	990021546	3865860	8335.071	1657	7625	7.625	0.317275	1.743823	39984.93
Oyster Bay	990021051	2362334	7604.329	1430	7375	7.375	0.304125	1.076400	23658.31
Oyster Bay	990021051	2362334	7604.329	1436	7125	7.125	0.290975	1.076400	22635.35

6 rows | 8-17 of 16 columns

3) A printed statement that says how much energy could be produced:

```
head(wind.tracts.full)

print(paste('The TX Gulf Coast GLO State Submerged Lands could produce up to: ', prettyNum(round(total_energy, 0),big.mark=",",scientific=FALSE), ' TWh per year.', sep = ''))
```

data.frame 6 x 16 | R Console

```
[1] "The TX Gulf Coast GLO State Submerged Lands could produce up to:  TWh per year."
```