Background & Purpose of Competition 3: Preparing for Real-Life Job Interviews

In your preparation for real-life interviews, you may face such a scenario that a company interviews you in groups. The company posts job descriptions for the business analyst(s) and invites you and other candidates to participate in a "group interview" project. You will be working with others, who usually are total strangers to you. The employer assigns technical or business tasks to the group, and each group member will present part of the group results in the end. The employer evaluates the overall performance and makes job offer(s) according to different performance levels. There are various possible outcomes from the interview process: (i) the team fails to deliver satisfactory results, and nobody is getting a job offer; (ii) only one offer is made to the best-performing person in the team; (iii) based on the company needs and the overall exceptional performance, all team members are getting individual job offers. Therefore, your team members are both your competitors and cooperators. Be prepared for a tough and challenging contribution to the project. One thing is clear: the whole team must deliver satisfactory results.

Prompt: Locational Power Grid Load Study

In the United States, multiple operators (known as ISO or Independent System Operators) provide electric power (or electricity) to regions and states.

- Example for regions: New England Area, the ISO is known as ISO-NE;
- Examples for states: in California, the ISO is known as CAISO, in New York State as NYISO.

The ISOs are responsible for providing a stable electricity supply and adjusting and balancing the supply with demand. The demand can vary due to different weather conditions such as heavy rain, severe heat, strong wind, heavy snow, cold snap, etc.

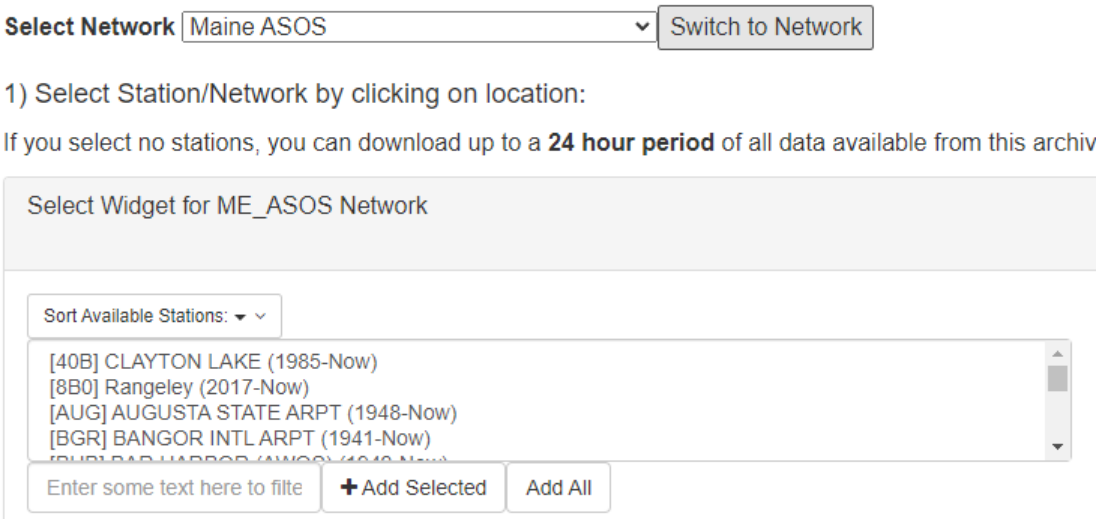
Therefore, it is imperative to predict the Actual Load of the power grid or get some approximations. If a high load is foreseen, the ISOs or local power grid operators under ISOs are turning on more generators to avoid power shortage; when a low load is foreseen, they turn off some generators, and there is no power over-supply.

For the Competition 3 Group Job Interview, we are focusing on the location of Central Maine, where the locational power grid is named Central Maine Power (CMP). We believe that Weather Condition is highly related to power grid Load. Therefore, we would like to use data science techniques to make predictions on the load.

Attached are the actual load data from 2019 to 2021 of CMP. The real load of each hour of every day is included in these excel files, each for one year. The data is also accessible from the link at the end of this problem description.

However, we need the historical weather data as well. Luckily, Iowa State University offers free downloaded historical weather data for you. You can visit the link below to download the weather data you want:
https://mesonet.agron.iastate.edu/request/download.phtml?network=MA_ASOS

Please kindly note that we need the **Average Weather Data** (e.g., temperature, humidity, etc.) from three cities for this project. Here are the names of the selected three cities of Maine: **Augusta, Waterville, and Portland**. You can find them easily by switching to Maine ASOS, as shown below:



After downloading the csv file, you may get confused with some column names. Here is some explanation:

Tmpf: Temperature in Fahrenheit
Dwpf: Dew Point in Fahrenheit
Relh: Relative humidity in Percentage
Sknt: Wind Speed in Knot
Mslp: Sea Level Pressure in Millibar
Vsby: Visibility in Miles
P01i: Precipitation Level in Inch
Feel: Temperature your body feel like in Fahrenheit

Targeted Outcomes:
Choose an **appropriate analytics model**, and use the data provided to make predictions for future Loads, i.e., predicting the hourly loads of CMP for **the next day**. Each hour's load should be included in your prediction, from 12:00 am, to 11:59 pm.

Use the following hints:
1. When choosing variables to build your model, please note that predictions of some weather metrics are hard to get. If you use those variables to build the model, it may cause a problem in your prediction.
2. Do you think Hour itself affects power grid load? Should it be included in the input variables?
3. Do you notice some "patterns" or "periodicity" of the load curve? Does your prediction have the same or similar pattern?

The judges realize that you are doing this with very limited time. Do the best you can. The answers you obtain may not be nice, neat, or cute – and that's okay. As you see them, present your findings and embrace the chance to hone your data wrangling, data merging, data visualization, and data modeling skills.

Deliverables, Part I: Your team will deliver slides and a 10-minute video presentation to answer Part I. The slides and your recording must be submitted before 9:00 AM on Saturday, March 26th.

Deliverables, Part II: During the final presentation of up to 15 minutes in length, you and your team may cover the points already included in your Part I recording, plus any additional insights from the data exploration or modeling process that may be of interest to the ABA community. Final presentations will be delivered live via Zoom on Saturday morning, beginning at 10:00 AM. EDT

CMP data source (2-18 – 2021):
https://www.cmpco.com/wps/portal/cmp/home/search/!ut/p/z1/zVRdc6IwFP0rrTP7SBM2SKFvKbW4jtpSi5q8MGlMNVtDEIN2__0G2of9qDB92mWYyWTuPTk3OfdcQMES0Jwd5JoZqXO2tXtC_Qy5k8HQi-A0mHghTDw0TuN4Hk-uXbBoSYC3oQtoO34OKKA8N4XZAMJVwdZlttFKfIF7wUq-qcN7XZVcRDo3IjdTvRKdnIAWXK4A8b-i5yD0hRM-I-54Tww5T0y4Th9yBv0A9X1PNAwNV3Zg28oe_jCdnd3oY77VbLX_JbyrRPnjj_CZwyqz0aXM10aoYsuMuLrq4ccsqSR_Gcv8pXcy5RYnp4P3pVa6ViH7ptha9Oo6LDms19d6XXRJU788PPHhv_B3sxGyeG_wOHRj5I4uu5R7w7cQdOLp7ykfVNB1B9pC0nQCsUVeZmESePGdTYj7cQCTWXQ99cI5imAfLA5SHEGa61LZdp_VJx65unhT_OJd8V31L1X_XCsvokmWpIMHe_NP-4a0-YD8Fy4gjQdI44AhBKOuLrMDTH7f7Si2U6Z-iFcDlh-NmUKlKkDvv5OmSyjvVUTsBmOc3YwxPj__CYI7q84!/?1dmy&mapping=%2Fcmp%2Fhome%2Fsearch&urile=wcm%3apath%3a%2FCMPAGR_SuppliersAndPartners%2FSuppliersandPartners%2FRegulatoryInformation%2FCMP_Transmission_Services%2FRNSDownloads