As part of our recruitment, we would like for you to demonstrate an involved but realistic problem you would face working on Hartree’s power desk. We would like for you to represent the implied volatility surface for natural gas (Henry Hub) European options listed on the CME exchange. Each day’s option quotes are listed publically at the link below and contain settlement prices for each option by strike for both the put and the call. You’ll also find a link to the futures settlement curve below. You’ll need to scrape this data in whatever programming language you feel most comfortable in (we would prefer python if you are well versed in it, but will not hold it against you if you’d prefer to do this in another language). You’ll also need to store this data locally so that it is consumable via the programming language you are using and can subsequently calculate and display the implied volatility surface in a useful way based on these option quotes. A simple curve/surface fit would do and for this exercise there is no need to implement a factor calibration process or to enforce no-arbitrage conditions.

We are purposefully leaving the instructions open-ended as we would like to see what you come up with on your own. That said, please do not hesitate to ask questions if you are stuck or require more guidance on a particular aspect of the project.

Futures link: <https://www.cmegroup.com/trading/energy/natural-gas/natural-gas_quotes_settlements_futures.html?optionProductId=1352&optionExpiration=1352-K0>

Options link: <https://www.cmegroup.com/trading/energy/natural-gas/natural-gas_quotes_settlements_options.html?optionProductId=1352&optionExpiration=1352-U9>

Please use the final settlement quotes and not the preliminary data as this can be less than reliable.

![A screenshot of a cell phone

Description automatically generated]()

1. Scrape option prices from CME website.  
2. Download prices and save to python pickle file. Downloaded options expiration calendar  
3. Attach underlying prices to the options prices by future maturity contract, attach expiration dates  
4. Calculate implied volatilities for each option in file, assuming a small interest rate for discounting  
5. Filter option prices based on cut-off of moneyness  
4. In the interest of time I decided to build the surface using calls only or puts only  
5. We save files to pickle files, the code checks to see if the source file exists, if it does not it goes to the CME ftp site  
 and downloads the csv file (had a lot of issues with scraping the actual site, code is included to show how I would  
 have approached the actual scraping, CME seems to have some restrictions around scraping their quote pages, some of the  
 python libraries were unable to get valid response (error code http 404)  
6. Utilized an OLS (Ordinary Least squares model provided by statsmodels python library to fit the vol surface to regression line  
7. Program prompts user for contract to evaluate e.g. 'NGN20'  
8. User can provide any contract maturity to generate vol curve.  
  
Improvements if I had more time:  
1. Add utilize a database server to create Tables (schema) to hold data  
2. Use both Puts and Calls to generate surface, possibly a polynomial based model, risk-reversals, straddles, strangles, etc.   
3. Create a surface using moneyness and maturity to show 3D structure and termstructure of volatility skew  
4. More error checking, unit testing  
5. etc.