Homework: CAS and PubChem Chemical Descriptors (Draft)

Due September 21, 2020

Please complete the following tasks; hopefully you'll be able to complete this homework assignment under an hour. Please see the accompanying *Handout: CAS and PubChem Chemical Descriptors* for notes on both the `R` coding and chemsitry needed/discussed in this assignement.

Tasks

On the Quercus course webpage:

1. Obtain your unique, individually assigned, chemical pollutant by complete the *Random Chemical Assignment* quiz on Quercus. This isn't worth any points, it'll simply assigne you a chemical.

Using R in R Studio:

2. Create a `data.frame` that includes your unique chemical and the six PFAS compounds in the `ExamplePFAS.csv` found on Quercus.
3. Using the `webchem` package, obtain the following molecular descriptors from the PubChem database:`MolecularFormula, MolecularWeight, CanonicalSMILES, IUPACName, XLogP, MonoisotopicMass, TPSA, Charge`, and `Volume3D`.
4. Combine both your compound `data.frame` and your obtained chemical descriptors into a single data.frame and save it as a comma-separated values (csv) file.
5. Draw your assigned chemical using it's SMILES; you can use PubChem Sketcher.
6. Create a plot of *molecular weight* vs. *LogP* of the seven compounds in your dataset. *You're plot must have at least one element differing from the example plot shown in the accompanying handout.*
    - Ideas to change your plot include applying a theme, modifying the axis labels and including a plot title), or anything else you can think of.
7. Discuss the results of your plot; example points of discussion are written in the accompanying handout.
8. **Submit the entirety of your work — code, plots, and discussion — as a PDF written in R markdown.** to Quercus.
    - See the *R Markdown tutorial* file on Quercus for additional support.
    - A skeleton R markdown file is provided on Quercus for your convenience, you'll need to fill it in with your own code and results.