**Data Science Test Widget Brain**

This assignment focuses on transshipments of a certain port. Each day, many vessels arrive in this port and are served by some stevedore(s). Four cargo types have been identified (ore, coal, oil, and petroleum), and vessels often carry a mixture of cargo types. For each unique vessel arrival (i.e. each row in the data), we would like a prediction of how much it transships (total of load & discharge activities) per cargo type. Variables of interest therefore are: discharge1, load1, discharge2, load2, discharge3, load3, discharge4 and load4.

The data for this case is stored in ‘VesselData.csv’ and contains historical data.

We would like you to provide us with a Jupyter notebook or Python script with the results of your endeavors, well enriched with comments elaborating on the steps taken, even if they did not lead you anywhere, and try to motivate your actions as much as possible. We would like to understand the approach you have taken and your line of thought.

**Explanation of variables in the data**

|  |  |
| --- | --- |
| Variable | Explanation |
| eta | Estimated time of arrival of vessel |
| ata | Actual time of arrival of vessel |
| atd | Actual time of departure of vessel |
| vesseldwt | Vessel deadweight tonnage |
| vesseltype | Vessel type |
| discharge[x] | Discharge amount of cargo type x |
| load[x] | Load amount of cargo type x |
| stevedorenames | (Anonymized) stevedore ID’s visited by vessel |
| hasnohamis | Boolean whether vessel has the HaMIS notification system |
| earliesteta | Estimated time of arrival of first entry to port |
| latesteta | Estimated time of arrival of last entry to port (vessel can spread transshipment(s) over multiple days) |
| traveltype | Travel type |
| previousportid | ID of previous port |
| nextportid | ID of next port |
| isremarkable | Boolean whether there is anything remarkable regarding the vessel |
| vesselid | Vessel ID |