These exercises are meant to make you explore how to build your own python pipeline classes . What you develop here , will be useful in working with future projects as well.You can attempt these problems in any order .

1

Develop a class which takes object type columns and coverts them pandas datetime type [given that these object type columns store dates]

Develop another class which can be used to extract weekday , month , day of month from date type columns and create cyclic features for these . consider cycle frequency to be 31 for day of month . [although frequency for day of month is not pprecise but it should not make much difference] .

Develop one more class which can take two date type columns and create a feature which has their difference .

Read data `Consumer_Complaints_train.csv` [Belongs to project 1] . Using classes created above ; create a pipeline for creating various features possible from the columns : `Date received` and `Date sent to company`

Note : Make use of class `VarSelector` [discussed in class and can be found in the shared class codes with name `mypipes.py` file] to filter these specific columns at the beginning of your pipeline component(s) .

2

Read file `Property_train.csv`. Break it into two parts : `train` and `test` using `train_test_split` function from `sklearn`.Identify categorical , numeric and date type columns in the data. Create a complete pipeline to prepare this data for predictive modeling . Fit this pipeline on train data and then use it to transform both train and test data.