|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57 | # -\*- coding: utf-8 -\*-  # modified from https://github.com/CamDavidsonPilon/lifelines/  <https://programtalk.com/vs2/python/1903/lifetimes/lifetimes/datasets/__init__.py/>  import pandas as pd  from .. import utils  from pkg\_resources import resource\_filename    \_\_all\_\_ = [      'load\_cdnow',      'load\_transaction\_data',      'load\_summary\_data\_with\_monetary\_value',      'load\_donations'  ]      def load\_dataset(filename, \*\*kwargs):      '''      Load a dataset from lifetimes.datasets        Parameters:      filename : for example "larynx.csv"      usecols : list of columns in file to use        Returns : Pandas dataframe      '''      return pd.read\_csv(resource\_filename('lifetimes', 'datasets/' + filename), \*\*kwargs)    def load\_donations(\*\*kwargs):      return load\_dataset('donations.csv', \*\*kwargs)    def load\_cdnow(\*\*kwargs):      return load\_dataset('cdnow\_customers.csv', \*\*kwargs)      def load\_transaction\_data(\*\*kwargs):      """      Returns a Pandas dataframe of transactional data. Looks like:                          date  id      0  2014-03-08 00:00:00   0      1  2014-05-21 00:00:00   1      2  2014-03-14 00:00:00   2      3  2014-04-09 00:00:00   2      4  2014-05-21 00:00:00   2        The data was artificially created using Lifetimes data generation routines. Data was generated      between 2014-01-01 to 2014-12-31.        """      return load\_dataset('example\_transactions.csv', \*\*kwargs)      def load\_summary\_data\_with\_monetary\_value(\*\*kwargs):      df = load\_dataset('cdnow\_customers\_transactions.csv', \*\*kwargs)      df.columns = ['customer\_id', 'frequency', 'recency', 'T', 'monetary\_value']      df = df.set\_index('customer\_id')      return df |