```
[ ]  import pandas as pd

     url = 'https://archive.ics.uci.edu/ml/machine-learning-databases/00382/c2k_data_comma.csv'
```

```
[ ]  df = pd.read_csv(url, index_col='nr')
```

UCI info: https://archive.ics.uci.edu/ml/datasets/Cargo+2000+Freight+Tracking+and+Tracing

```
Attribute Information:

nr - unique id for process instance of overall process - domain: [1â€¦3942]

i1_legid - unique id across all transport legs (note: also to 'empty' legs are assigned an id) of incoming tran
i1_rcs_p - planned duration (minutes) of incoming transport leg 1 (RCS: Freight Check in) - domain: [LONGINT]
i1_rcs_e - effective (i.e., actual) duration (minutes) of incoming transport leg 1 (RCS: Freight Check in) - do
i1_dep_1_p - planned duration (minutes) of incoming transport leg 1 (DEP: Departure Segment 1) - domain: [LONGI
i1_dep_1_e - effective (i.e., actual) duration (minutes) of incoming transport leg 1 (DEP: Departure Segment 1)
i1_dep_1_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i1_rcf_1_p - planned duration (minutes) of incoming transport leg 1 (RCF: Arrival Segment 1) - domain: [LONGINT
i1_rcf_1_e - effective (i.e., actual) duration (minutes) of incoming transport leg 1 (RCF: Arrival Segment 1) -
i1_rcf_1_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i1_dep_2_p - planned duration (minutes) of incoming transport leg 1 (DEP: Departure Segment 2) - domain: [LONGI
i1_dep_2_e - effective (i.e., actual) duration (minutes) of incoming transport leg 1 (DEP: Departure Segment 2)
i1_dep_2_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i1_rcf_2_p - planned duration (minutes) of incoming transport leg 1 (RCF: Arrival Segment 2) - domain: [LONGINT
i1_rcf_2_e - effective (i.e., actual) duration (minutes) of incoming transport leg 1 (RCF: Arrival Segment 2) -
i1_rcf_2_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i1_dep_3_p - planned duration (minutes) of incoming transport leg 1 (DEP: Departure Segment 3) - domain: [LONGI
i1_dep_3_e - effective (i.e., actual) duration (minutes) of incoming transport leg 1 (DEP: Departure Segment 3)
i1_dep_3_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i1_rcf_3_p - planned duration (minutes) of incoming transport leg 1 (RCF: Arrival Segment 3) - domain: [LONGINT
i1_rcf_3_e - effective (i.e., actual) duration (minutes) of incoming transport leg 1 (RCF: Arrival Segment 3) -
i1_rcf_3_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i1_dlv_p - planned duration (minutes) of incoming transport leg 1 (DLV: Freight Delivery) - domain: [LONGINT]
i1_dlv_e - effective (i.e., actual) duration (minutes) of incoming transport leg 1 (DLV: Freight Delivery) - do
i1_hops - number of segments (hops) in the transport leg of incoming transport leg 1 - domain: [1..4]

i2_legid - unique id across all transport legs (note: also to 'empty' legs are assigned an id) of incoming tran
i2_rcs_p - planned duration (minutes) of incoming transport leg 2 (RCS: Freight Check in) - domain: [LONGINT]
i2_rcs_e - effective (i.e., actual) duration (minutes) of incoming transport leg 2 (RCS: Freight Check in) - do
i2_dep_1_p - planned duration (minutes) of incoming transport leg 2 (DEP: Departure Segment 1) - domain: [LONGI
i2_dep_1_e - effective (i.e., actual) duration (minutes) of incoming transport leg 2 (DEP: Departure Segment 1)
i2_dep_1_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i2_rcf_1_p - planned duration (minutes) of incoming transport leg 2 (RCF: Arrival Segment 1) - domain: [LONGINT
i2_rcf_1_e - effective (i.e., actual) duration (minutes) of incoming transport leg 2 (RCF: Arrival Segment 1) -
i2_rcf_1_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i2_dep_2_p - planned duration (minutes) of incoming transport leg 2 (DEP: Departure Segment 2) - domain: [LONGI
i2_dep_2_e - effective (i.e., actual) duration (minutes) of incoming transport leg 2 (DEP: Departure Segment 2)
i2_dep_2_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i2_rcf_2_p - planned duration (minutes) of incoming transport leg 2 (RCF: Arrival Segment 2) - domain: [LONGINT
i2_rcf_2_e - effective (i.e., actual) duration (minutes) of incoming transport leg 2 (RCF: Arrival Segment 2) -
i2_rcf_2_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i2_dep_3_p - planned duration (minutes) of incoming transport leg 2 (DEP: Departure Segment 3) - domain: [LONGI
i2_dep_3_e - effective (i.e., actual) duration (minutes) of incoming transport leg 2 (DEP: Departure Segment 3)
i2_dep_3_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i2_rcf_3_p - planned duration (minutes) of incoming transport leg 2 (RCF: Arrival Segment 3) - domain: [LONGINT
i2_rcf_3_e - effective (i.e., actual) duration (minutes) of incoming transport leg 2 (RCF: Arrival Segment 3) -
i2_rcf_3_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i2_dlv_p - planned duration (minutes) of incoming transport leg 2 (DLV: Freight Delivery) - domain: [LONGINT]
i2_dlv_e - effective (i.e., actual) duration (minutes) of incoming transport leg 2 (DLV: Freight Delivery) - do
i2_hops - number of segments (hops) in the transport leg of incoming transport leg 2 - domain: [1..4]

i3_legid - unique id across all transport legs (note: also to 'empty' legs are assigned an id) of incoming tran
i3_rcs_p - planned duration (minutes) of incoming transport leg 3 (RCS: Freight Check in) - domain: [LONGINT]
i3_rcs_e - effective (i.e., actual) duration (minutes) of incoming transport leg 3 (RCS: Freight Check in) - do
i3_dep_1_p - planned duration (minutes) of incoming transport leg 3 (DEP: Departure Segment 1) - domain: [LONGI
i3_dep_1_e - effective (i.e., actual) duration (minutes) of incoming transport leg 3 (DEP: Departure Segment 1)
i3_dep_1_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i3_rcf_1_p - planned duration (minutes) of incoming transport leg 3 (RCF: Arrival Segment 1) - domain: [LONGINT
i3_rcf_1_e - effective (i.e., actual) duration (minutes) of incoming transport leg 3 (RCF: Arrival Segment 1) -
i3_rcf_1_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
```

```
i3_dep_2_p - planned duration (minutes) of incoming transport leg 3 (DEP: Departure Segment 2) - domain: [LONGI
i3_dep_2_e - effective (i.e., actual) duration (minutes) of incoming transport leg 3 (DEP: Departure Segment 2)
i3_dep_2_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i3_rcf_2_p - planned duration (minutes) of incoming transport leg 3 (RCF: Arrival Segment 2) - domain: [LONGINT
i3_rcf_2_e - effective (i.e., actual) duration (minutes) of incoming transport leg 3 (RCF: Arrival Segment 2) -
i3_rcf_2_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i3_dep_3_p - planned duration (minutes) of incoming transport leg 3 (DEP: Departure Segment 3) - domain: [LONGI
i3_dep_3_e - effective (i.e., actual) duration (minutes) of incoming transport leg 3 (DEP: Departure Segment 3)
i3_dep_3_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i3_rcf_3_p - planned duration (minutes) of incoming transport leg 3 (RCF: Arrival Segment 3) - domain: [LONGINT
i3_rcf_3_e - effective (i.e., actual) duration (minutes) of incoming transport leg 3 (RCF: Arrival Segment 3) -
i3_rcf_3_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of
i3_dlv_p - planned duration (minutes) of incoming transport leg 3 (DLV: Freight Delivery) - domain: [LONGINT]
i3_dlv_e - effective (i.e., actual) duration (minutes) of incoming transport leg 3 (DLV: Freight Delivery) - dor
i3_hops - number of segments (hops) in the transport leg of incoming transport leg 3 - domain: [1..4]

o_legid - unique id across all transport legs (note: also to 'empty' legs are assigned an id) of outgoing transp
o_rcs_p - planned duration (minutes) of outgoing transport leg (RCS: Freight Check in) - domain: [LONGINT]
o_rcs_e - effective (i.e., actual) duration (minutes) of outgoing transport leg (RCS: Freight Check in) - domai
o_dep_1_p - planned duration (minutes) of outgoing transport leg (DEP: Departure Segment 1) - domain: [LONGINT]
o_dep_1_e - effective (i.e., actual) duration (minutes) of outgoing transport leg (DEP: Departure Segment 1) - (
o_dep_1_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of (
o_rcf_1_p - planned duration (minutes) of outgoing transport leg (RCF: Arrival Segment 1) - domain: [LONGINT]
o_rcf_1_e - effective (i.e., actual) duration (minutes) of outgoing transport leg (RCF: Arrival Segment 1) - dor
o_rcf_1_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of (
o_dep_2_p - planned duration (minutes) of outgoing transport leg (DEP: Departure Segment 2) - domain: [LONGINT]
o_dep_2_e - effective (i.e., actual) duration (minutes) of outgoing transport leg (DEP: Departure Segment 2) - (
o_dep_2_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of (
o_rcf_2_p - planned duration (minutes) of outgoing transport leg (RCF: Arrival Segment 2) - domain: [LONGINT]
o_rcf_2_e - effective (i.e., actual) duration (minutes) of outgoing transport leg (RCF: Arrival Segment 2) - dor
o_rcf_2_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of (
o_dep_3_p - planned duration (minutes) of outgoing transport leg (DEP: Departure Segment 3) - domain: [LONGINT]
o_dep_3_e - effective (i.e., actual) duration (minutes) of outgoing transport leg (DEP: Departure Segment 3) - (
o_dep_3_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of (
o_rcf_3_p - planned duration (minutes) of outgoing transport leg (RCF: Arrival Segment 3) - domain: [LONGINT]
o_rcf_3_e - effective (i.e., actual) duration (minutes) of outgoing transport leg (RCF: Arrival Segment 3) - dor
o_rcf_3_place - unique id for airport (original IATA codes have been masked due to confidentiality reasons) of (
o_dlv_p - planned duration (minutes) of outgoing transport leg (DLV: Freight Delivery) - domain: [LONGINT]
o_dlv_e - effective (i.e., actual) duration (minutes) of outgoing transport leg (DLV: Freight Delivery) - domai
o_hops - number of segments (hops) in the transport leg of outgoing transport leg - domain: [1..4]

legs - number of incoming transport legs of overall process - domain: [1..3]
```

```
[ ]  df.tail()
```

```
[ ]  #  Que forma tiene nuestro dataset?
     df.shape
```

```
[ ]  #  Que forma tiene nuestro dataset? Despues de nan
     df.dropna().shape
```

```
[ ]  # Reemplazar los signos de interrogacion
     df = df.replace(to_replace='?', value=0, inplace=True)
```

```
[ ]  # casting
     df = df.astype('float64')
```

```
[ ]  df.dtypes
```

```
[ ]  df.head()
```

```
[ ]
```