

Best Submission

✓ Successful

Submitted by Team Brute Force Everything 22 days ago

Private Score

1.00000

Public Score

1.00000

A Big Thanks to the notebook contributor to let me get the ideas to solve this question as a beginner!

1. 3 hooman 1 cat - Logistics

2. week5

Tips to get 1.0

There are some tips to get 1.0 from this discussion. Another big thanks!<https://www.kaggle.com/c/open-shopee-code-league-logistic/discussion/165972>:

1. You can easily count the weekday difference with the help of numpy.busday\_count function! Check the "weekmask" and "holidays" in numpy document.

2. We use pd.to\_datetime to process the epoch time from "deliveryordersmarch.csv". However, don't forget we need to convert the time to GMT+8

[1]. <https://www.kaggle.com/c/open-shopee-code-league-logistic/discussion/165829>

## Set environment

```
In [1]: import numpy as np
import pandas as pd
```

## Data manipulate

```
In [2]: data = pd.read_csv('../input/logistics-shopee-code-league/delivery_orders_march.csv')
```

```
In [3]: # look at data
data.head()
```

	orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress
0	2215676524	1583138397	1.583385e+09	NaN	Baging Idl BUENAVISTA,PATAG,CAGAYAN Buagsong,c...	Pantranco vill. 417 Warehouse katipunan 532 L...
1	2219624609	1583309968	1.583463e+09	1.583799e+09	coloma's quzom CASANAS Site1 Masiyan 533A Stol...	BLDG 210A Moras C42l 2816,168 church Complex ...
...	...	...	...	...	...	...
3	2221066352	1583419016	1.583556e+09	NaN	616Espritu MARTINVILLE,MANUYO #5paraiso kengli...	999mall 201,26 Vitaruel Barretto gent number...
4	2222478803	1583318305	1.583480e+09	NaN	L042 Summerbreeze1 L2[Balanay analyn Lot760 C...	OGGMANIL Hiyas Finest MAYSILO nagdiwang Lt4

```
In [4]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3176313 entries, 0 to 3176312
Data columns (total 6 columns):
#   Column      Dtype
---  ---
0   orderid     int64
1   pick        int64
2   1st_deliver_attempt float64
3   2nd_deliver_attempt float64
4   buyeraddress object
5   selleraddress object
dtypes: float64(2), int64(2), object(2)
memory usage: 145.4+ MB
```

```
In [5]: # get the state of the address by getting the last str in the address
```

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress
0	2215676524	1583138397	1.583385e+09	NaN	manila
1	2219624609	1583309968	1.583463e+09	1.583799e+09	manila
2	2220979489	1583306434	1.583460e+09	NaN	manila
3	2221066352	1583419016	1.583556e+09	NaN	manila
4	2222478803	1583318305	1.583480e+09	NaN	luzon

```
In [6]: # calculate the business days to deliver before late
temp = []
for i, j in data[['buyeraddress', 'selleraddress']].itertuples(index=False):
    if (i == 'manila' and j == 'manila'):
        temp.append(3)
    elif (
        (i == 'manila' and j == 'luzon')
        or (i == 'luzon' and j == 'manila')
        or (i == 'luzon' and j == 'luzon')
    ):
        temp.append(5)
    else:
        temp.append(7)
data['days'] = temp
data.head()
```

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days
0	2215676524	1583138397	1.583385e+09	NaN	manila	3
1	2219624609	1583309968	1.583463e+09	1.583799e+09	manila	3
2	2220979489	1583306434	1.583460e+09	NaN	manila	3
3	2221066352	1583419016	1.583556e+09	NaN	manila	3
4	2222478803	1583318305	1.583480e+09	NaN	luzon	5

```
In [7]: # change unix time to date
# need to convert to gmt+8
data[['pick', '1st_deliver_attempt', '2nd_deliver_attempt']] += 8*60*60
data['pick'] = pd.to_datetime(data['pick'], unit='s').dt.date
data['1st_deliver_attempt'] = pd.to_datetime(data['1st_deliver_attempt'], unit='s').dt.date
data['2nd_deliver_attempt'] = data['2nd_deliver_attempt'].replace(np.nan, 0) # change nan to 0 or else can't be process
data['2nd_deliver_attempt'] = pd.to_datetime(data['2nd_deliver_attempt'], unit='s').dt.date
data.head()
```

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5

```
In [8]: # count how many days of business day taken for the 1pick and 2pick
holiday = ['2020-03-08', '2020-03-25', '2020-03-30', '2020-03-31']
data['1st_pick'] = np.busday_count(data['pick'], data['1st_deliver_attempt'], weekmask='1111110', holidays=holiday)
data['2nd_pick'] = np.busday_count(data['1st_deliver_attempt', data['2nd_deliver_attempt']
```

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	-15708	
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3	2	-15709
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3	2	-15710
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5	2	-15709

```
In [9]: # check if is late
data['is_late'] = (data['1st_pick'] > data['days']) | (data['2nd_pick'] > 3)
data.head()
```

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	3	-15708
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3	2	-15709
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3	2	-15710
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5	2	-15709

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	3	-15708
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3	2	-15709
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3	2	-15710
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5	2	-15709

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	3	-15708
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3	2	-15709
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3	2	-15710
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5	2	-15709

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	3	-15708
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3	2	-15709
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3	2	-15710
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5	2	-15709

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	3	-15708
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3	2	-15709
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3	2	-15710
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5	2	-15709

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	3	-15708
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3	2	-15709
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3	2	-15710
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5	2	-15709

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	3	-15708
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3	2	-15709
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3	2	-15710
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5	2	-15709

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	3	-15708
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3	2	-15709
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3	2	-15710
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5	2	-15709

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	3	-15708
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3	2	-15709
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3	2	-15710
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5	2	-15709

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	3	-15708
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3	2	-15709
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3	2	-15710
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5	2	-15709

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	3	-15708
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3	2	-15709
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3	2	-15710
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5	2	-15709

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	3	-15708
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	2220979489	2020-03-04	2020-03-06	1970-01-01	manila	manila	3	2	-15709
3	2221066352	2020-03-05	2020-03-07	1970-01-01	manila	manila	3	2	-15710
4	2222478803	2020-03-04	2020-03-06	1970-01-01	luzon	manila	5	2	-15709

orderid	pick	1st_deliver_attempt	2nd_deliver_attempt	buyeraddress	selleraddress	days	1st_pick	2nd_pick	
0	2215676524	2020-03-02	2020-03-05	1970-01-01	manila	manila	3	3	-15708
1	2219624609	2020-03-04	2020-03-06	2020-03-10	manila	manila	3	2	3
2	222								