

# **ML-MINOR-APRIL**

**INFORMATION ON SUPPLY CHAINS  
USED BY DATACO GLOBAL**

**SUMMER INTERNSHIP MINOR  
PROJECT**

**DONE BY: ABHISHEK CHAKRABORTY**

**achk31@gmail.com**

**MACHINE LEARNING APRIL BATCH 2021**

## QUESTION 1:

List all the cities and the respective number of pending orders in ‘India’?

Sol: The cities of India mentioned in the data set are: Bikaner, Jabalpur, Bangalore, Pune, Jaipur, Ranchi, Durg, Patna, Faridabad, Fatehpur, Kota, Guwahati, Rohtak, Jamnagar, Thane, Dewas, Kulti, Imphal, Gorakhpur, Lucknow, Hubli, Nagpur, Ajmer, Amravati, Raipur, Tiruneveli, Darbhanga, Aligarh, Bhopal, Saharanpur, Vadodra, Delhi, Puducherry, Kanpur, Jabalpur, Moradabad, Belgaum, Hyderabad, Nellore & Korba.

Based on the information below:

```
In [13]: data2.info()

<class 'pandas.core.frame.DataFrame'>
Int64Index: 79 entries, 1 to 62424
Data columns (total 2 columns):
 #   Column      Non-Null Count   Dtype  
--- 
 0   Order City    79 non-null     object 
 1   Order Status  79 non-null     object 
dtypes: object(2)
memory usage: 1.9+ KB
```

There are total 79 pending orders in India.

## QUESTION 2:

Which country has the most suspected fraud?

Sol: Based on the info mentioned below:

```
In [19]: fraud_data.value_counts()
```

```
Out[19]: Order Country    Order Status
Estados Unidos  SUSPECTED_FRAUD    109
Francia        SUSPECTED_FRAUD    70
México         SUSPECTED_FRAUD    56
Brasil          SUSPECTED_FRAUD    45
Alemania        SUSPECTED_FRAUD    32
...
Noruega         SUSPECTED_FRAUD    1
Pakistán        SUSPECTED_FRAUD    1
Sudán           SUSPECTED_FRAUD    1
Suecia          SUSPECTED_FRAUD    1
Afganistán      SUSPECTED_FRAUD    1
Length: 68, dtype: int64
```

It appears that the United States has the most suspected fraud, followed by France & Mexico.

## QUESTION 3:

Which product(s) has/have the maximum discount?

Sol: Based on the info mentioned below:

```
In [24]: maxdisc.value_counts()
Out[24]: Order Item Discount  Product Name
100.000000  Field & Stream Sportsman 16 Gun Fire Safe    148
124.989998  Nike Men's Free 5.0+ Running Shoe      78
113.010002  Web Camera                      22
255.000000  Dell Laptop                     17
240.000000  Dell Laptop                     17
270.000000  Dell Laptop                     16
135.000000  Dell Laptop                     15
300.000000  Dell Laptop                     12
115.370003  Porcelain crafts                 12
150.000000  Dell Laptop                     12
133.149994  Lawn mower                       11
105.000000  Dell Laptop                     11
180.000000  Dell Laptop                     11
225.000000  Dell Laptop                     9
375.000000  Dell Laptop                     9
195.000000  Dell Laptop                     8
106.519997  Lawn mower                       7
100.000000  LIJA Women's Mid-Length Panel Golf Shorts   2
108.000000  Bowflex SelectTech 1090 Dumbbells       1
100.000000  GoPro HERO3+ Black Edition Camera     1
          SOLE E35 Elliptical                   1
          Garmin Forerunner 910XT GPS Watch      1
500.000000  SOLE E35 Elliptical                   1
112.489998  Brooks Women's Ghost 6 Running Shoe      1
120.000000  Bowflex SelectTech 1090 Dumbbells       1
123.750000  The North Face Women's Recon Backpack   1
124.989998  TaylorMade White Smoke IN-12 Putter      1
400.000000  SOLE E35 Elliptical                   1
200.000000  SOLE E35 Elliptical                   1
240.000000  SOLE E35 Elliptical                   1
125.000000  LIJA Women's Mid-Length Panel Golf Shorts   1
dtype: int64
```

It appears that products like Field & Stream Sportsman 16 Gun Fire Safe Dell Laptop has the maximum discount.

## QUESTION 4:

How many successfully processed orders have days for shipping (real) > days for shipment (scheduled)?

Sol: Based on the information mentioned below:

```
In [195]: successdata2.info()
<class 'pandas.core.frame.DataFrame'>
Int64Index: 15960 entries, 6 to 65533
Data columns (total 3 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Days for shipping (real)    15960 non-null   int64  
 1   Days for shipment (scheduled) 15960 non-null   int64  
 2   Order Status               15960 non-null   object 
dtypes: int64(2), object(1)
memory usage: 498.8+ KB
```

There are total 15960 processed orders for days for shipping (real) > days for shipment (scheduled).