```
import numpy as np
```

```
l1 = dbutils.fs.ls('dbfs:/FileStore/demo-files')
len(l1)
```

```
Out[169]: 3
```

```
from pyspark.sql.functions import *
```

```
df_main = spark.createDataFrame([], schema= StructType([]))
```

```
for file in l1:
    print(f'\nFile Path :- {file.path}')
    df = spark.read.csv(file.path, header=True)
#     arr = [data[0]for data in df.select('address').collect()]
    df_bool = df.withColumn('address', col('address').rlike('[a-zA-Z][*!@^&?%#$]').alias('address'))
    df_bool = df_bool.na.fill(False, subset=['address'])
    display(df_bool)
    myarr = [data[0] for data in df_bool.select('address').collect()]
    print(f'Address columns condition check:- {myarr}')
    signal = np.all(myarr)
    print(f'Signal to keep File:- {signal}')
    if signal == True:
        df_main = df
```

```
File Path :- dbfs:/FileStore/demo-files/sample1.csv
```

Table

	name	address
1	ravi	true
2	suraj	true
3	sumit	true

3 rows

```
Address columns condition check:- [True, True, True]
Signal to keep File:- True

File Path :- dbfs:/FileStore/demo-files/sample2.csv
```

Table

	name	address
1	ravi	false
2	suraj	false
3	sumit	false
4	anuj	false

4 rows

```
Address columns condition check:- [False, False, False, False]
Signal to keep File:- False

File Path :- dbfs:/FileStore/demo-files/sample3.csv
```

Table

	name	address
1	ravi	true
2	suraj	true
3	sumit	true
4	vijay	false
5	james	true

5 rows

```
Address columns condition check:- [True, True, True, False, True]
Signal to keep File:- False
```

```
display(df_main)
```

Table

	name	address
1	ravi	Pune#
2	suraj	Pune?
3	sumit	Pune*

3 rows