```
#    02 - Ecommerce Purchases
##   August 3, 2022
#    1 Pandas

# ** Import pandas and read in the Ecommerce Purchases csv file. **
import pandas as pd

# Read Ecommerce_Purchases.txt
ecom_pur= pd.read_csv(r"C:\Users\Administrator\Downloads\
Ecommerce_Purchases.txt")
ecom_pur

                                                 Address    Lot AM or
PM  \
0     16629 Pace Camp Apt. 448\r\nAlexisborough, NE ...  46 in
PM
1     9374 Jasmine Spurs Suite 508\r\nSouth John, TN...  28 rn
PM
2                     Unit 0065 Box 5052\r\nDPO AP 27450  94 vE
PM
3               7780 Julia Fords\r\nNew Stacy, WA 45798  36 vm
PM
4     23012 Munoz Drive Suite 337\r\nNew Cynthia, TX...  20 IE
AM
...                                                 ...    ...      ..
.
9995    966 Castaneda Locks\r\nWest Juliafurt, CO 96415  92 XI
PM
9996  832 Curtis Dam Suite 785\r\nNorth Edwardburgh,...  41 JY
AM
9997            Unit 4434 Box 6343\r\nDPO AE 28026-0283  74 Zh
AM
9998             0096 English Rest\r\nRoystad, IA 12457  74 cL
PM
9999   40674 Barrett Stravenue\r\nGrimesville, WI 79682  64 Hr
AM

                                           Browser Info  \
0     Opera/9.56.(X11; Linux x86_64; sl-SI) Presto/2...
1     Opera/8.93.(Windows 98; Win 9x 4.90; en-US) Pr...
2     Mozilla/5.0 (compatible; MSIE 9.0; Windows NT ...
3     Mozilla/5.0 (Macintosh; Intel Mac OS X 10_8_0 ...
4     Opera/9.58.(X11; Linux x86_64; it-IT) Presto/2...
...                                                 ...
9995  Mozilla/5.0 (Windows NT 5.1) AppleWebKit/5352 ...
9996  Mozilla/5.0 (compatible; MSIE 9.0; Windows NT ...
9997  Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_7...
9998  Mozilla/5.0 (Macintosh; Intel Mac OS X 10_8_8;...
9999  Mozilla/5.0 (X11; Linux i686; rv:1.9.5.20) Gec...

                              Company       Credit Card CC Exp Date  \
```

```
0                    Martinez-Herman  6011929061123406       02/20
1     Fletcher, Richards and Whitaker  3337758169645356       11/18
2          Simpson, Williams and Pham      675957666125       08/19
3     Williams, Marshall and Buchanan  6011578504430710       02/24
4           Brown, Watson and Andrews  6011456623207998       10/25
...                               ...               ...         ...
9995                   Randall-Sloan   342945015358701       03/22
9996         Hale, Collins and Wilson   210033169205009       07/25
9997                   Anderson Ltd  6011539787356311       05/21
9998                       Cook Inc   180003348082930       11/17
9999                     Greene Inc  4139972901927273       02/19

      CC Security Code                 CC Provider  \
0                  900                JCB 16 digit
1                  561                  Mastercard
2                  699                JCB 16 digit
3                  384                   Discover
4                  678  Diners Club / Carte Blanche
...                ...                         ...
9995               838                JCB 15 digit
9996               207                JCB 16 digit
9997                 1               VISA 16 digit
9998               987            American Express
9999               302                JCB 15 digit

                             Email
Job  \
0                 pdunlap@yahoo.com  Scientist, product/process
development
1                anthony41@reed.com                      Drilling
engineer
2     amymiller@morales-harrison.com              Customer service
manager
3        brent16@olson-robinson.info                      Drilling
engineer
4        christopherwright@gmail.com                           Fine
artist
...                              ...
...
9995          iscott@wade-garner.com
Printmaker
9996              mary85@hotmail.com                        Energy
engineer
9997               tyler16@gmail.com                    Veterinary
surgeon
9998         elizabethmoore@reid.net               Local government
officer
9999           rachelford@vaughn.com                 Embryologist,
clinical
```

```
           IP Address Language  Purchase Price
0     149.146.147.205       el           98.14
1        15.160.41.51       fr           70.73
2      132.207.160.22       de            0.95
3        30.250.74.19       es           78.04
4        24.140.33.94       es           77.82
...               ...      ...             ...
9995    29.73.197.114       it           82.21
9996   121.133.168.51       pt           25.63
9997    156.210.0.254       el           83.98
9998     55.78.26.143       es           38.84
9999  176.119.198.199       el           67.59

[10000 rows x 14 columns]
```

```
#Check the head of the DataFrame.
ecom_pur.head()
```

```
                                              Address    Lot AM or
PM  \
0  16629 Pace Camp Apt. 448\r\nAlexisborough, NE ...  46 in       PM

1  9374 Jasmine Spurs Suite 508\r\nSouth John, TN...  28 rn       PM

2                Unit 0065 Box 5052\r\nDPO AP 27450  94 vE       PM

3           7780 Julia Fords\r\nNew Stacy, WA 45798  36 vm       PM

4  23012 Munoz Drive Suite 337\r\nNew Cynthia, TX...  20 IE       AM


                                       Browser Info  \
0  Opera/9.56.(X11; Linux x86_64; sl-SI) Presto/2...
1  Opera/8.93.(Windows 98; Win 9x 4.90; en-US) Pr...
2  Mozilla/5.0 (compatible; MSIE 9.0; Windows NT ...
3  Mozilla/5.0 (Macintosh; Intel Mac OS X 10_8_0 ...
4  Opera/9.58.(X11; Linux x86_64; it-IT) Presto/2...

                          Company       Credit Card CC Exp Date  \
0                 Martinez-Herman  6011929061123406       02/20
1  Fletcher, Richards and Whitaker  3337758169645356       11/18
2       Simpson, Williams and Pham     675957666125       08/19
3  Williams, Marshall and Buchanan  6011578504430710       02/24
4        Brown, Watson and Andrews  6011456623207998       10/25

   CC Security Code                CC Provider  \
0               900                JCB 16 digit
1               561                  Mastercard
2               699                JCB 16 digit
3               384                    Discover
```

```
4               678  Diners Club / Carte Blanche

                            Email
Job  \
0               pdunlap@yahoo.com  Scientist, product/process
development
1              anthony41@reed.com                       Drilling
engineer
2  amymiller@morales-harrison.com                Customer service
manager
3     brent16@olson-robinson.info                       Drilling
engineer
4     christopherwright@gmail.com                              Fine
artist

        IP Address Language  Purchase Price
0  149.146.147.205       el           98.14
1     15.160.41.51       fr           70.73
2   132.207.160.22       de            0.95
3     30.250.74.19       es           78.04
4     24.140.33.94       es           77.82
```

```
# ** How many rows and columns are there? **
ecom_pur.info()

# computing number of rows
rows = len(ecom_pur.axes[0])

# computing number of columns
cols = len(ecom_pur.axes[1])

print("Number of Rows: ", rows)
print("Number of Columns: ", cols)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 14 columns):
Address             10000 non-null object
Lot                 10000 non-null object
AM or PM            10000 non-null object
Browser Info        10000 non-null object
Company             10000 non-null object
Credit Card         10000 non-null int64
CC Exp Date         10000 non-null object
CC Security Code    10000 non-null int64
CC Provider         10000 non-null object
Email               10000 non-null object
Job                 10000 non-null object
IP Address          10000 non-null object
Language            10000 non-null object
```

```
Purchase Price      10000 non-null float64
dtypes: float64(1), int64(2), object(11)
memory usage: 1.1+ MB
Number of Rows:  10000
Number of Columns:  14
```

```
# ** What is the average Purchase Price? **
ecom_pur["Purchase Price"].mean()
```

```
50.34730200000025
```

```
#  ** What were the highest and lowest purchase prices? **
ecom_pur["Purchase Price"].max()
```

```
99.99
```

```
ecom_pur["Purchase Price"].min()
```

```
0.0
```

```
# ** How many people have English 'en' as their Language of choice on
the website? **
ecom_pur[ecom_pur["Language"]=="en"].count()
```

```
Address             1098
Lot                 1098
AM or PM            1098
Browser Info        1098
Company             1098
Credit Card         1098
CC Exp Date         1098
CC Security Code    1098
CC Provider         1098
Email               1098
Job                 1098
IP Address          1098
Language            1098
Purchase Price      1098
dtype: int64
```

```
# ** How many people have the job title of "Lawyer" ? **
ecom_pur[ecom_pur["Job"]=="Lawyer"].info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 30 entries, 470 to 9979
Data columns (total 14 columns):
Address             30 non-null object
Lot                 30 non-null object
AM or PM            30 non-null object
Browser Info        30 non-null object
Company             30 non-null object
Credit Card         30 non-null int64
```

```
CC Exp Date         30 non-null object
CC Security Code    30 non-null int64
CC Provider         30 non-null object
Email               30 non-null object
Job                 30 non-null object
IP Address          30 non-null object
Language            30 non-null object
Purchase Price      30 non-null float64
dtypes: float64(1), int64(2), object(11)
memory usage: 3.5+ KB
```

```
#** How many people made the purchase during the AM and how many
people made the purchase during PM ? **
ecom_pur['AM or PM'].value_counts()
```

```
PM    5068
AM    4932
Name: AM or PM, dtype: int64
```

```
# ** What are the 5 most common Job Titles? **
ecom_pur['Job'].value_counts().head(5)
```

```
Interior and spatial designer        31
Lawyer                               30
Social researcher                    28
Purchasing manager                   27
Research officer, political party    27
Name: Job, dtype: int64
```

```
# ** Someone made a purchase that came from Lot: “90 WT” , what was
the Purchase Price for this transaction? **
ecom_pur[ecom_pur['Lot']=='90 WT']['Purchase Price']
```

```
513    75.1
Name: Purchase Price, dtype: float64
```

```
# ** What is the email of the person with the following Credit Card
Number: 4926535242672853 **
ecom_pur[ecom_pur['Credit Card']==4926535242672853]['Email']
```

```
1234    bondellen@williams-garza.com
Name: Email, dtype: object
```

```
# ** How many people have American Express as their Credit Card
Provider and made a purchase above $95 ?**
ecom_pur[(ecom_pur['CC Provider']=='American Express') &
(ecom_pur['Purchase Price']>95)].count()
```

```
Address             39
Lot                 39
AM or PM            39
Browser Info        39
```

```
Company             39
Credit Card         39
CC Exp Date         39
CC Security Code    39
CC Provider         39
Email               39
Job                 39
IP Address          39
Language            39
Purchase Price      39
dtype: int64
```

```
# ** Hard: How many people have a credit card that expires in 2025? **
sum(ecom_pur['CC Exp Date'].apply(lambda x: x[3:]) == '25')
```

```
1033
```

```
# ** Hard: What are the top 5 most popular email providers/hosts (e.g.
gmail.com, yahoo.com, etc…)**
ecom_pur['Email'].apply(lambda x: x.split('@')
[1]).value_counts().head(5)
```

```
hotmail.com     1638
yahoo.com       1616
gmail.com       1605
smith.com         42
williams.com      37
Name: Email, dtype: int64
```