

Import Pandas and read in the Ecommerce Purchases csv file set it to a DataFrame called ecom.

```
import pandas as pd
```

[1] ✓ 3.9s

Python

```
ecom = pd.read_csv(r"data/Ecommerce Purchases")
```

[2] ✓ 0.3s

Python

Check the head of the Dataframe.

```
ecom.head()
```

[3] ✓ 0.1s

Python

...

	Address	Lot	AM or PM	Browser Info	Company	Credit Card	CC Exp Date	CC Security Code	CC Provider	Email	
0	16629 Pace Camp Apt. 448\nAlexisborough,	46 in	PM	Opera/9.56. (X11; Linux x86_64; sl-	Martinez- Herman	6011929061123406	02/20	900	JCB 16 digit	pdunlap@yahoo.com	product

How many rows and columns are there?

```
len(ecom.columns)
```

[4] ✓ 0.1s

... 14

```
len(ecom.index)
```

[5] ✓ 0.9s

... 10000

```
ecom.info()
```

[6] ✓ 0.1s

... <class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 14 columns):
Column Non-Null Count Dtype
--- -

```
ecom.describe()
```

[7] ✓ 0.1s

...

	Credit Card	CC Security Code	Purchase Price
count	1.000000e+04	10000.000000	10000.000000
mean	2.341374e+15	907.217800	50.347302
std	2.256103e+15	1589.693035	29.015836
min	6.040186e+10	0.000000	0.000000
25%	3.056322e+13	280.000000	25.150000
50%	8.699942e+14	548.000000	50.505000
75%	4.492298e+15	816.000000	75.770000
max	6.012000e+15	9993.000000	99.990000

What is the average Purchases Price?

```
ecom['Purchase Price'].mean()
```

[8] ✓ 0.6s

...

50.347302

What is the highest and lowest purchase price?

```
ecom['Purchase Price'].max()
```

[9] ✓ 0.9s

... 99.99

```
ecom['Purchase Price'].min()
```

[10] ✓ 0.9s

... 0.0

How many people have English 'en' as their Language of choice on the website?

```
# ecom[ecom['Language'] == 'en'].count()
```

```
ecom[ecom['Language'] == 'en']['Language'].count()
```

```
# len(ecom[ecom['Language']=="en"].index)
```

How Many People have the job title of "Lawyer"?

```
# ecom[ecom['Job'] == "Lawyer"].count()

ecom[ecom["Job"] == 'Lawyer']['Job'].count()

# len(ecom[ecom['Job']=="Lawyer"].index)
```

[12] ✓ 0.7s

... 30

How many people made the purchases during the AM and how many people made the purchase during PM?

```
ecom['AM or PM'].value_counts()
```

[13] ✓ 0.7s

... PM 5068

AM 4932

Name: AM or PM, dtype: int64

What are the 5 most common Job Titles?



```
ecom['Job'].value_counts().head()
```

[14]

✓ 0.8s

```
... Interior and spatial designer    31
     Lawyer                        30
     Social researcher               28
     Purchasing manager             27
     Designer, jewellery            27
     Name: Job, dtype: int64
```

Someone made a purchase that came from lot: "90 WT", what was the Purchase price for this transacton?

```
ecom[ecom["Lot"] == "90 WT"]['Purchase Price']
```

[15]

✓ 0.7s

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

What is the email of the person with the following Credit Card Number: 4926535242672853

```
ecom[ecom["Credit Card"] == 4926535242672853]['Email']
```

6] ✓ 0.1s

• 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[(ecom["CC Provider"] == "American express") & (ecom["Purchase Price"] > 95)].index
```

7] ✓ 0.9s

• Int64Index([], dtype='int64')

How many People have a credit card that expires in 2025?

```
# sum(ecom["CC Exp Date"].apply(lambda exp: exp[3:] == '25'))  
ecom[ecom["CC Exp Date"].apply(lambda exp: exp[3:] == '25')].count()
```

[32] ✓ 0.1s

```
... Address      1033  
    Lot          1033  
    AM or PM     1033  
    Browser Info 1033  
    Company      1033  
    Credit Card   1033  
    CC Exp Date   1033  
    CC Security Code 1033  
    CC Provider   1033  
    Email         1033  
    Job           1033  
    IP Address    1033  
    Language      1033  
    Purchase Price 1033  
    dtype: int64
```


What are the top 5 most popular email providers



```
ecom["Email"].apply(lambda email: email.split("@")[1]).value_counts().head(5)
```

[42]

✓ 0.8s

...

hotmail.com 1638

yahoo.com 1616

gmail.com 1605

smith.com 42

williams.com 37

Name: Email, dtype: int64