

ASSIGNMENT 11

KARTIK THAKUR

How many rows and columns are there?

```
In [32]: import pandas as pd
import numpy as np
file="Ecommerce Purchases"
df = pd.read_csv(file)
copy=df.copy()
copy
```

Out[32]:

AM or PM	Browser Info	Company	Credit Card	CC Exp Date	CC Security Code	CC Provider	
PM	Opera/9.56.(X11; Linux x86_64; sl- SI) Presto/2...	Martinez- Herman	6011929061123406	02/20	900	JCB 16 digit	pdunlap@ya
PM	Opera/8.93. (Windows 98; Win 9x 4.90; en-US) Pr...	Fletcher, Richards and Whitaker	3337758169645356	11/18	561	Mastercard	anthony41@r
PM	Mozilla/5.0 (compatible; MSIE 9.0; Windows NT ...	Simpson, Williams and Pham	675957666125	08/19	699	JCB 16 digit	amymiller@ harr
PM	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_8_0 ...	Williams, Marshall and Buchanan	6011578504430710	02/24	384	Discover	brent16@olson-robi
AM	Opera/9.58.(X11; Linux x86_64; it- IT) Presto/2...	Brown, Watson and Andrews	6011456623207998	10/25	678	Diners Club / Carte Blanche	christopherwright@g
...
PM	Mozilla/5.0 (Windows NT 5.1) AppleWebKit/5352 ...	Randall- Sloan	342945015358701	03/22	838	JCB 15 digit	iscott@wade-ga
AM	Mozilla/5.0 (compatible; MSIE 9.0; Windows NT ...	Hale, Collins and Wilson	210033169205009	07/25	207	JCB 16 digit	mary85@hot
AM	Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_7...	Anderson Ltd	6011539787356311	05/21	1	VISA 16 digit	tyler16@g
PM	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_8_8;...	Cook Inc	180003348082930	11/17	987	American Express	elizabethmoore@
AM	Mozilla/5.0 (X11; Linux i686; rv:1.9.5.20) Gec...	Greene Inc	4139972901927273	02/19	302	JCB 15 digit	rachelford@vau



In [33]: df.shape # .shape finds the shape of the dataframe

Out[33]: (10000, 14)

2. What is the average Purchase Price?

```
In [34]: copy["Purchase Price"].mean() # gives average purchase price
```

```
Out[34]: 50.34730200000025
```

4. How many people have French 'fr' as their Language of choice on the website?

```
In [35]: # Count the number of people who speak English
fr_speakers = (df['Language'].str.contains('fr')).sum()
fr_speakers
```

```
Out[35]: 1097
```

5. How many people have the job title of "Drilling engineer" ?

```
In [36]: (df['Job'].str.contains('Drilling engineer')).sum()
```

```
Out[36]: 21
```

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

```
In [37]: (df['AM or PM'].str.contains('AM')).sum()
```

```
Out[37]: 4932
```

```
In [38]: (df['AM or PM'].str.contains('PM')).sum()
```

```
Out[38]: 5068
```

7. What are the 5 most common Job Titles?

```
In [39]: copy['Job'].value_counts().head()
```

```
Out[39]: Interior and spatial designer    31
         Lawyer                          30
         Social researcher                 28
         Purchasing manager               27
         Designer, jewellery              27
         Name: Job, dtype: int64
```

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

```
In [40]: row = df.loc[df["Credit Card"] == 4926535242672853]
row
```

Out[40]:

Address	Lot	AM or PM	Browser Info	Company	Credit Card	CC Exp Date	CC Security Code	CC Provider
70 Maria Manors Suite 707 Joneshaven, MN ...	82 UX	PM	Mozilla/5.0 (X11; Linux x86_64; rv:1.9.5.20) G...	Cole, King and Bowers	4926535242672853	09/21	188	American Express

```
In [41]: email = row["Email"].values[0]
email
```

Out[41]: 'bondellen@williams-garza.com'

9. How many people have American Express as their Credit Card Provider and made a purchase above \$45 ?

```
In [44]: num_people = len(df[(df['Credit Card'] == 'American Express') & (df['Purchase'] > 45)])
num_people
```

Out[44]: 0

10. How many people have a credit card that expires in 2025?

```
In [57]: df['Year'] = df['CC Exp Date'].str[-2:]
expiring_in_2025 = df[df['Year'] == '25']
count=len(expiring_in_2025)
count
```

Out[57]: 1033