

# ecommerce-purchases-script

February 4, 2024

```
[ ]: import pandas as pd
```

```
[2]: data = pd.read_csv('D:/data analytics/Python for data analytics/Python_
↳Projects/Ecommerce Purchase/Ecommerce Purchases')
```

## 1.Display Top 10 Rows of Data

```
[5]: data.head(10)
```

```
[5]:
```

	Address	Lot	AM or PM	\
0	16629 Pace Camp Apt. 448\nAlexisborough, NE 77...	46 in	PM	
1	9374 Jasmine Spurs Suite 508\nSouth John, TN 8...	28 rn	PM	
2	Unit 0065 Box 5052\nDPO AP 27450	94 vE	PM	
3	7780 Julia Fords\nNew Stacy, WA 45798	36 vm	PM	
4	23012 Munoz Drive Suite 337\nNew Cynthia, TX 5...	20 IE	AM	
5	7502 Powell Mission Apt. 768\nTravisland, VA 3...	21 XT	PM	
6	93971 Conway Causeway\nAndersonburgh, AZ 75107	96 Xt	AM	
7	260 Rachel Plains Suite 366\nCastroberg, WV 24...	96 pG	PM	
8	2129 Dylan Burg\nNew Michelle, ME 28650	45 JN	PM	
9	3795 Dawson Extensions\nLake Tinafort, ID 88739	15 Ug	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...	
5	Mozilla/5.0 (Macintosh; U; PPC Mac OS X 10_8_5...	
6	Mozilla/5.0 (compatible; MSIE 7.0; Windows NT ...	
7	Mozilla/5.0 (X11; Linux i686) AppleWebKit/5350...	
8	Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_7...	
9	Mozilla/5.0 (X11; Linux i686; rv:1.9.7.20) 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
5	Silva-Anderson	30246185196287	07/25
6	Gibson and Sons	6011398782655569	07/24
7	Marshall-Collins	561252141909	06/25
8	Galloway and Sons	180041795790001	04/24
9	Rivera, Buchanan and Ramirez	4396283918371	01/17

	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
5	7169	Discover
6	714	VISA 16 digit
7	256	VISA 13 digit
8	899	JCB 16 digit
9	931	American Express

	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
5	ynguyen@gmail.com	Fish farm manager
6	olivia04@yahoo.com	Dancer
7	phillip48@parks.info	Event organiser
8	kdavis@rasmussen.com	Financial manager
9	qcoleman@hunt-huerta.com	Forensic scientist

	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
5	55.96.152.147	ru	25.15
6	127.252.144.18	de	88.56
7	224.247.97.150	pt	44.25
8	146.234.201.229	ru	59.54
9	236.198.199.8	zh	95.63

## 2. Check Last 10 Rows of the Dataset

```
[6]: data.tail(10)
```

[6]:

	Address	Lot	AM or PM	\
9990	75731 Molly Springs\nWest Danielle, VT 96934-5102	93 ty	PM	
9991	PSC 8165, Box 8498\nAPO AP 60327-0346	50 dA	AM	
9992	885 Allen Mountains Apt. 230\nWallhaven, LA 16995	40 vH	PM	
9993	7555 Larson Locks Suite 229\nEllisburgh, MA 34...	72 jg	PM	
9994	6276 Rojas Hollow\nLake Louis, WY 56410-7837	93 Ex	PM	
9995	966 Castaneda Locks\nWest Juliafurt, CO 96415	92 XI	PM	
9996	832 Curtis Dam Suite 785\nNorth Edwardburgh, T...	41 JY	AM	
9997	Unit 4434 Box 6343\nDPO AE 28026-0283	74 Zh	AM	
9998	0096 English Rest\nRoystad, IA 12457	74 cL	PM	
9999	40674 Barrett Stravenue\nGrimesville, WI 79682	64 Hr	AM	

	Browser Info	\
9990	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_7_4;...	
9991	Mozilla/5.0 (compatible; MSIE 8.0; Windows NT ...	
9992	Mozilla/5.0 (Macintosh; PPC Mac OS X 10_6_5) A...	
9993	Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_8...	
9994	Opera/9.68.(X11; Linux x86_64; sl-SI) 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	\
9990	Pace, Vazquez and Richards	869968197049750	04/24	
9991	Snyder Inc	4221582137197481	02/24	
9992	Wells Ltd	4664825258997302	10/20	
9993	Colon and Sons	30025560104631	10/25	
9994	Ritter-Smith	3112186784121077	01/25	
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	Email	\
9990	877	JCB 15 digit	andersonmichael@sherman.biz	
9991	969	Voyager	kking@wise-liu.com	
9992	431	Discover	bberry@wright.net	
9993	629	Maestro	chelseawilliams@lopez.biz	
9994	1823	Maestro	iroberts@gmail.com	
9995	838	JCB 15 digit	iscott@wade-garner.com	
9996	207	JCB 16 digit	mary85@hotmail.com	
9997	1	VISA 16 digit	tyler16@gmail.com	
9998	987	American Express	elizabethmoore@reid.net	
9999	302	JCB 15 digit	rachelford@vaughn.com	

	Job	IP Address	Language	Purchase Price
9990	Early years teacher	54.170.3.185	ru	18.35
9991	IT sales professional	254.25.31.156	el	25.93
9992	Set designer	174.173.51.32	de	67.96
9993	Designer, exhibition/display	177.46.82.128	el	65.61
9994	Education officer, museum	242.44.112.18	zh	31.85
9995	Printmaker	29.73.197.114	it	82.21
9996	Energy engineer	121.133.168.51	pt	25.63
9997	Veterinary surgeon	156.210.0.254	el	83.98
9998	Local government officer	55.78.26.143	es	38.84
9999	Embryologist, clinical	176.119.198.199	el	67.59

### 3.Check Datatype of Each Column.

```
[7]: data.dtypes
```

```
[7]: Address          object
     Lot              object
     AM or PM         object
     Browser Info     object
     Company          object
     Credit Card      int64
     CC Exp Date      object
     CC Security Code int64
     CC Provider      object
     Email            object
     Job              object
     IP Address       object
     Language         object
     Purchase Price   float64
     dtype: object
```

### 4.Check Null Values in the Dataset

```
[8]: data.isnull().sum()
```

```
[8]: Address          0
     Lot              0
     AM or PM         0
     Browser Info     0
     Company          0
     Credit Card      0
     CC Exp Date      0
     CC Security Code 0
     CC Provider      0
     Email            0
     Job              0
```

```
IP Address          0
Language            0
Purchase Price      0
dtype: int64
```

## 5.How many Rows and Columns are there on our Dataset?

```
[9]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 14 columns):
 #   Column                Non-Null Count  Dtype
---  -
 0   Address               10000 non-null  object
 1   Lot                   10000 non-null  object
 2   AM or PM              10000 non-null  object
 3   Browser Info          10000 non-null  object
 4   Company               10000 non-null  object
 5   Credit Card           10000 non-null  int64
 6   CC Exp Date           10000 non-null  object
 7   CC Security Code      10000 non-null  int64
 8   CC Provider           10000 non-null  object
 9   Email                 10000 non-null  object
10   Job                   10000 non-null  object
11   IP Address            10000 non-null  object
12   Language              10000 non-null  object
13   Purchase Price        10000 non-null  float64
dtypes: float64(1), int64(2), object(11)
memory usage: 1.1+ MB
```

## 6.Highest and Lowest Purchase Prices

```
[10]: data['Purchase Price'].max()
```

```
[10]: 99.99
```

```
[11]: data['Purchase Price'].min()
```

```
[11]: 0.0
```

## 7.Average Purchase Price

```
[12]: data['Purchase Price'].mean()
```

```
[12]: 50.347302
```

## 8.How many People have French 'fr' as their language?

```
[13]: len(data[data['Language'] == 'fr'])
```

```
[13]: 1097
```

### 9.Job title Contains Engineer

```
[14]: len(data[data['Job'].str.contains('engineer', case=False)])
```

```
[14]: 984
```

### 10.Find Email of the person with the following IP address:132.207.160.22

```
[15]: data[data['IP Address'] == '132.207.160.22']
```

```
[15]:
```

	Address	Lot	AM or PM	\
2	Unit 0065 Box 5052\ndPO AP 27450	94	vE	PM

  

	Browser Info	\
2	Mozilla/5.0 (compatible; MSIE 9.0; Windows NT ...	

  

	Company	Credit Card	CC Exp Date	CC Security Code	\
2	Simpson, Williams and Pham	675957666125	08/19	699	

  

	CC Provider	Email	Job	\
2	JCB 16 digit	amymiller@morales-harrison.com	Customer service manager	

  

	IP Address	Language	Purchase Price
2	132.207.160.22	de	0.95

### 11.How many people have Mastercard as their credit card provider and made a purchase above 50 ?

```
[18]: data[(data['CC Provider'] == 'Mastercard') & (data['Purchase Price'] > 50)].  
      ↪count()
```

```
[18]: Address          405  
      Lot            405  
      AM or PM       405  
      Browser Info   405  
      Company        405  
      Credit Card    405  
      CC Exp Date    405  
      CC Security Code 405  
      CC Provider    405  
      Email          405  
      Job            405  
      IP Address     405  
      Language       405
```

```
Purchase Price      405
dtype: int64
```

**12. Find Email of the person with the following credit card number: 4664825258997302**

```
[19]: data[data['Credit Card'] == 4664825258997302]['Email']
```

```
[19]: 9992      bberry@wright.net
      Name: Email, dtype: object
```

**13. How many People Purchase during the AM and how many people purchase during PM?**

```
[20]: data['AM or PM'].value_counts()
```

```
[20]: AM or PM
      PM      5068
      AM      4932
      Name: count, dtype: int64
```

**14. How many people have a credit card that expires in 2020?**

```
[23]: len(data[data['CC Exp Date'].apply(lambda x: x[3:] == '20')])
```

```
[23]: 988
```

**15. Top 5 Most Popular Email providers (eg: gmail.com, yahoo.com, etc...)**

```
[25]: data['Email'].apply(lambda x: x.split('@')[1]).value_counts().head()
```

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