"QUESTIONS"

- 1. Display Top 10 Rows of The Dataset
- 2. Check Last 10 Rows of The Dataset
- 3. Check Datatype of Each Column
- 4. Check null values in the dataset
- 5. How many rows and columns are there in our Dataset?
- 6. Highest and Lowest Purchase Prices.
- 7. Average Purchase Price
- 8. How many people have French 'fr' as their Language?
- 9. Job Title Contains Engineer
- 10. Find The Email of the person with the following IP Address: 132.207.160.22
- 11. How many People have Mastercard as their Credit Card Provider and made a purchase above 50?
- 12. Find the email of the person with the following Credit Card Number: 4664825258997302
- 13. How many people purchase during the AM and how many people purchase during PM?
- 14. How many people have a credit card that expires in 2020?
- 15. What are the top 5 most popular email providers (e.g. gmail.com, yahoo.com, etc...)

```
from google.colab import files
files.upload()

Choose Files kaggle.json
• kaggle.json(application/json) - 65 bytes, last modified: 3/19/2025 - 100% done
Saving kaggle.json to kaggle.json
{'kaggle.json': b'{"username":"onkar0629","key":"d3b74c8466360314c093d84f478041ec"}'}

import os
os.makedirs("/root/.kaggle", exist_ok=True)
!mv kaggle.json /root/.kaggle/
!chmod 600 /root/.kaggle/kaggle.json
```

!kaggle datasets download -d utkarsharya/ecommerce-purchases

Warning: Looks like you're using an outdated API Version, please consider updating (serv Dataset URL: https://www.kaggle.com/datasets/utkarsharya/ecommerce-purchases
License(s): unknown ecommerce-purchases.zip: Skipping, found more recently modified local copy (use --force

```
import zipfile
with zipfile.ZipFile("ecommerce-purchases.zip", "r") as zip_ref:
    zip ref.extractall("ecommerce data")
import pandas as pd
df = pd.read csv("ecommerce data/Ecommerce Purchases")
print(df.head()) # Display the first 5 rows
\rightarrow
                                                   Address
                                                               Lot AM or PM
        16629 Pace Camp Apt. 448\nAlexisborough, NE 77...
                                                            46 in
                                                                         PM
       9374 Jasmine Spurs Suite 508\nSouth John, TN 8...
                                                             28 rn
                                                                         PM
                         Unit 0065 Box 5052\nDPO AP 27450
     2
                                                            94 vE
                                                                         PM
     3
                    7780 Julia Fords\nNew Stacy, WA 45798
                                                                         PM
                                                            36 vm
        23012 Munoz Drive Suite 337\nNew Cynthia, TX 5...
                                                             20 IE
                                                                         AΜ
                                              Browser Info
     0 Opera/9.56.(X11; Linux x86_64; s1-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...
                                               Credit Card CC Exp Date \
                                 Company
     0
                        Martinez-Herman
                                          6011929061123406
                                                                  02/20
        Fletcher, Richards and Whitaker
     1
                                                                  11/18
                                          3337758169645356
     2
             Simpson, Williams and Pham
                                              675957666125
                                                                  08/19
       Williams, Marshall and Buchanan 6011578504430710
     3
                                                                  02/24
              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
                                         Scientist, product/process development
                     pdunlap@yahoo.com
     1
                    anthony41@reed.com
                                                               Drilling engineer
     2
        amymiller@morales-harrison.com
                                                       Customer service manager
     3
           brent16@olson-robinson.info
                                                               Drilling engineer
           christopherwright@gmail.com
     4
                                                                     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
```

import pandas as pd

30.250.74.19

24.140.33.94

3

es

es

78.04

77.82

data = pd.read_csv("ecommerce_data/Ecommerce Purchases")
data



	Address	Lot	AM or PM	Browser Info	Company	Credit Card	CC Exp Date	
0	16629 Pace Camp Apt. 448\nAlexisborough, NE 77	46 in	РМ	Opera/9.56.(X11; Linux x86_64; sl- SI) Presto/2	Martinez- Herman	6011929061123406	02/20	
1	9374 Jasmine Spurs Suite 508\nSouth John, TN 8	28 rn	PM	Opera/8.93. (Windows 98; Win 9x 4.90; en-US) Pr	Fletcher, Richards and Whitaker	3337758169645356	11/18	
2	Unit 0065 Box 5052\nDPO AP 27450	94 vE	PM	Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 	Simpson, Williams and Pham	675957666125	08/19	
3	7780 Julia Fords\nNew Stacy, WA 45798	36 vm	PM	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_8_0 	Williams, Marshall and Buchanan	6011578504430710	02/24	
4	23012 Munoz Drive Suite 337\nNew Cynthia, TX 5	20 IE	AM	Opera/9.58.(X11; Linux x86_64; it- IT) Presto/2	Brown, Watson and Andrews	6011456623207998	10/25	
9995	966 Castaneda Locks\nWest Juliafurt, CO 96415	92 XI	PM	Mozilla/5.0 (Windows NT 5.1) AppleWebKit/5352 	Randall- Sloan	342945015358701	03/22	
9996	832 Curtis Dam Suite 785\nNorth Edwardburgh, T	41 JY	AM	Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 	Hale, Collins and Wilson	210033169205009	07/25	
9997	Unit 4434 Box 6343\nDPO AE 28026- 0283	74 Zh	AM	Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_7	Anderson Ltd	6011539787356311	05/21	
9998	0096 English Rest\nRoystad, IA 12457	74 cL	PM	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_8_8;	Cook Inc	180003348082930	11/17	
9999	40674 Barrett Stravenue\nGrimesville, WI 79682	64 Hr	AM	Mozilla/5.0 (X11; Linux i686; rv:1.9.5.20) Gec	Greene Inc	4139972901927273	02/19	
10000 rows × 14 columns								

Next steps: (

Generate code with data

View recommended plots

New interactive sheet

#1. Display Top 10 Rows of The Dataset
data.head(10)

-		_
	→	$\overline{}$
×	Ť	_

	Address	Lot	AM or PM	Browser Info	Company	Credit Card	C Ex Dat
0	16629 Pace Camp Apt. 448\nAlexisborough, NE 77	46 in	PM	Opera/9.56.(X11; Linux x86_64; sI-SI) Presto/2	Martinez- Herman	6011929061123406	02/2
1	9374 Jasmine Spurs Suite 508\nSouth John, TN 8	28 rn	РМ	Opera/8.93. (Windows 98; Win 9x 4.90; en-US) Pr	Fletcher, Richards and Whitaker	3337758169645356	11/1
2	Unit 0065 Box 5052\nDPO AP 27450	94 vE	PM	Mozilla/5.0 (compatible; MSIE 9.0; Windows NT	Simpson, Williams and Pham	675957666125	08/1
3	7780 Julia Fords\nNew Stacy, WA 45798	36 vm	PM	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_8_0 	Williams, Marshall and Buchanan	6011578504430710	02/2
4	23012 Munoz Drive Suite 337\nNew Cynthia, TX 5	20 IE	AM	Opera/9.58.(X11; Linux x86_64; it-IT) Presto/2	Brown, Watson and Andrews	6011456623207998	10/2
5	7502 Powell Mission Apt. 768\nTravisland, VA 3	21 XT	РМ	Mozilla/5.0 (Macintosh; U; PPC Mac OS X 10_8_5	Silva- Anderson	30246185196287	07/2
6	93971 Conway Causeway\nAndersonburgh, AZ 75107	96 Xt	AM	Mozilla/5.0 (compatible; MSIE 7.0; Windows NT	Gibson and Sons	6011398782655569	07/2
7	260 Rachel Plains Suite 366\nCastroberg, WV 24	96 pG	PM	Mozilla/5.0 (X11; Linux i686) AppleWebKit/5350	Marshall- Collins	561252141909	06/2
8	2129 Dylan Burg\nNew Michelle, ME 28650	45 JN	PM	Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_7	Galloway and Sons	180041795790001	04/2
9	3795 Dawson Extensions\nLake Tinafort, ID 88739	15 Ug	AM	Mozilla/5.0 (X11; Linux i686; rv:1.9.7.20) Gec	Rivera, Buchanan and Ramirez	4396283918371	01/1
4							

Next steps: (Generate code with data

View recommended plots

New interactive sheet

#2 Check Last 10 Rows of The Dataset
data.tail(10)



	Address	Lot	AM or PM	Browser Info	Company	Credit Card	CC Exp Date
9990	75731 Molly Springs\nWest Danielle, VT 96934- 5102	93 ty	PM	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_7_4;	Pace, Vazquez and Richards	869968197049750	04/24
9991	PSC 8165, Box 8498\nAPO AP 60327- 0346	50 dA	AM	Mozilla/5.0 (compatible; MSIE 8.0; Windows NT 	Snyder Inc	4221582137197481	02/24
9992	885 Allen Mountains Apt. 230\nWallhaven, LA 16995	40 vH	PM	Mozilla/5.0 (Macintosh; PPC Mac OS X 10_6_5) A	Wells Ltd	4664825258997302	10/20
9993	7555 Larson Locks Suite 229\nEllisburgh, MA 34	72 jg	PM	Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_8	Colon and Sons	30025560104631	10/25
9994	6276 Rojas Hollow\nLake Louis, WY 56410-7837	93 Ex	PM	Opera/9.68.(X11; Linux x86_64; sl- SI) Presto/2	Ritter- Smith	3112186784121077	01/25
9995	966 Castaneda Locks\nWest Juliafurt, CO 96415	92 XI	PM	Mozilla/5.0 (Windows NT 5.1) AppleWebKit/5352 	Randall- Sloan	342945015358701	03/22
9996	832 Curtis Dam Suite 785\nNorth Edwardburgh, T	41 JY	AM	Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 	Hale, Collins and Wilson	210033169205009	07/25
9997	Unit 4434 Box 6343\nDPO AE 28026- 0283	74 Zh	AM	Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_7	Anderson Ltd	6011539787356311	05/21
9998	0096 English Rest\nRoystad, IA 12457	74 cL	PM	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_8_8;	Cook Inc	180003348082930	11/17
9999	40674 Barrett Stravenue\nGrimesville, WI 79682	64 Hr	AM	Mozilla/5.0 (X11; Linux i686; rv:1.9.5.20) Gec	Greene Inc	4139972901927273	02/19
4							•

#3 Check Datatype of Each Column
data.dtypes



	0
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

dtype: object

Purchase Price

#4 Check null values in the dataset
data.isnull().sum()

float64



0
0
0
0
0
0
0
0
0
0
0
0
0
0
0

dtype: int64

#5 How many rows and columns are there in our Dataset?
len(data.columns)

→ 14

len(data)

→ 10000

data.info()

<<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 14 columns):

20.00		0010111110	
#	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

```
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
     10 Job
                           10000 non-null object
     11 IP Address
                           10000 non-null object
     12 Language
                         10000 non-null object
                           10000 non-null float64
     13 Purchase Price
    dtypes: float64(1), int64(2), object(11)
    memory usage: 1.1+ MB
#6 Highest and Lowest Purchase Prices.
data.columns
→ Index(['Address', 'Lot', 'AM or PM', 'Browser Info', 'Company', 'Credit Card',
            'CC Exp Date', 'CC Security Code', 'CC Provider', 'Email', 'Job',
            'IP Address', 'Language', 'Purchase Price'],
          dtvpe='object')
data['Purchase Price'].max()
→ 99.99
data['Purchase Price'].min()
→ 0.0
#7 Average Purchase Price
data['Purchase Price'].mean()
→ np.float64(50.347302)
#8 How many people have French 'fr' as their Language?
data.columns
→ Index(['Address', 'Lot', 'AM or PM', 'Browser Info', 'Company', 'Credit Card',
            'CC Exp Date', 'CC Security Code', 'CC Provider', 'Email', 'Job',
            'IP Address', 'Language', 'Purchase Price'],
          dtype='object')
data[data['Language']=='fr']
```



	Address	Lot	AM or PM	Browser Info	Company	Credit Card	CC Exp Date	Sec
1	9374 Jasmine Spurs Suite 508\nSouth John, TN 8	28 rn	PM	Opera/8.93. (Windows 98; Win 9x 4.90; en-US) Pr	Fletcher, Richards and Whitaker	3337758169645356	11/18	
19	125 Hall Summit\nBoothton, IL 41721	99 CU	РМ	Mozilla/5.0 (compatible; MSIE 7.0; Windows NT	Turner- Mckinney	676343504830	02/20	
53	PSC 9431, Box 7059\nAPO AA 29285-1363	14 qD	АМ	Opera/9.34. (X11; Linux x86_64; it-IT) Presto/2	Higgins, Cardenas and Kennedy	869972604798355	08/17	
76	49206 Campbell Port\nNorth Cliffordshire, HI 3	71 iu	PM	Mozilla/5.0 (Macintosh; U; PPC Mac OS X 10_5_9	Jacobs- Tucker	6011343518820988	01/17	
82	493 Smith Valleys Suite 004\nNew Madelineville	35 Is	РМ	Mozilla/5.0 (iPod; U; CPU iPhone OS 4_0 like M	Robinson, Johnston and Valdez	4351359627548412	06/17	
9941	43757 Brown Lodge\nValerieberg, PR 88518	68 DG	AM	Mozilla/5.0 (compatible; MSIE 5.0; Windows NT	Jones, Williams and Dunn	6011508474487291	04/26	
9947	32827 Carl Inlet\nSouth Nicole, NY 38081-6636	40 Dp	РМ	Mozilla/5.0 (Windows; U; Windows 95) AppleWebK	Pruitt- Flores	3158564208754951	03/17	
9951	518 Brown Dam\nGarciaside, IN 33057	71 wq	АМ	Mozilla/5.0 (compatible; MSIE 8.0; Windows NT	Crosby, Vasquez and Ballard	5256758441931287	12/17	
9977	02182 Keith Expressway\nEast Shannon, CT 20578	34 RL	АМ	Mozilla/5.0 (compatible; MSIE 9.0; Windows NT	Deleon, Jacobson and Benton	4186094003664688	06/21	

Opera/8.58.

(Windows NT 5.0; it-IT) Frye 6011944199756993 12/24 Fresto/2.9....

1097 rows × 14 columns

len(data[data['Language']=='fr'])

→ 1097

data[data['Language']=='fr'].count()

÷		_
Ľ	4	÷
Ξ	7	

	0
Address	1097
Lot	1097
AM or PM	1097
Browser Info	1097
Company	1097
Credit Card	1097
CC Exp Date	1097
CC Security Code	1097
CC Provider	1097
Email	1097
Job	1097
IP Address	1097
Language	1097
Purchase Price	1097

dtype: int64

#9 Job Title Contains Engineer
data.columns

data[data['Job'].str.contains('engineer',case=False)]



	Address	Lot	AM or PM	Browser Info	Company	Credit Card	CC Exp Date	Se
1	9374 Jasmine Spurs Suite 508\nSouth John, TN 8	28 rn	PM	Opera/8.93. (Windows 98; Win 9x 4.90; en- US) Pr	Fletcher, Richards and Whitaker	3337758169645356	11/18	
3	7780 Julia Fords\nNew Stacy, WA 45798	36 vm	PM	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_8_0	Williams, Marshall and Buchanan	6011578504430710	02/24	
50	41159 Michael Centers\nAdamsfort, RI 37108-6674	46 Ce	PM	Mozilla/5.0 (Windows 98; Win 9x 4.90; sl-SI; r	Wright, Williams and Mendez	4008586485908075	05/19	
55	27635 Maureen Bypass Apt. 883\nSandraview, SD 	59 LJ	AM	Mozilla/5.0 (iPod; U; CPU iPhone OS 3_3 like M	Sims- Lyons	3158113629128344	09/19	
60	7126 Katherine Squares\nPerkinsview, CO 97299	63 qu	AM	Opera/8.68. (X11; Linux x86_64; en- US) Presto/2	Marshall- Fernandez	349767747049645	05/20	
9948	95544 Johnson Isle Suite 939\nMichaelberg, RI 	91 bW	AM	Opera/8.36. (X11; Linux x86_64; sl- SI) Presto/2	Fox- Peterson	4762924304307	03/17	
9952	9991 Vaughn Hills\nRacheltown, PA 55409	36 KC	PM	Mozilla/5.0 (X11; Linux i686; rv:1.9.5.20) Gec	Ward, Smith and Castillo	6011679271321726	09/19	
9970	0060 Keith Stream\nWestport, CO 47097	11 nt	PM	Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_8	Carpenter, Good and Hart	6011485664704662	07/19	
9977	02182 Keith Expressway\nEast Shannon CT 20578	34 RL	AM	Mozilla/5.0 (compatible; MSIE 9.0;	Deleon, Jacobson and	4186094003664688	06/21	

```
יייסווווטוו, טו בטטוט...
                                                V V II IUU VV 3
                                                               Benton
                                                   NT ...
                                              Mozilla/5.0
                                                                 Hale,
         832 Curtis Dam Suite
                                             (compatible;
                                                               Collins
                                  41
 9996
                    785\nNorth
                                       AM
                                               MSIE 9.0;
                                                                          210033169205009 07/25
                                  JY
                                                                  and
                                                Windows
             Edwardburgh, T...
                                                               Wilson
                                                   NT ...
984 rows × 14 columns
```

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

 \rightarrow 984

#10 Find The Email of the person with the following IP Address: 132.207.160.22 data.columns

→ Index(['Address', 'Lot', 'AM or PM', 'Browser Info', 'Company', 'Credit Card', 'CC Exp Date', 'CC Security Code', 'CC Provider', 'Email', 'Job', 'IP Address', 'Language', 'Purchase Price'], dtype='object')

data[data['IP Address']=="132.207.160.22"]['Email']

 $\overline{\Sigma}$

Email

2 amymiller@morales-harrison.com

dtype: object

#11 How many People have Mastercard as their Credit Card Provider and made a purchase above data.columns

```
Index(['Address', 'Lot', 'AM or PM', 'Browser Info', 'Company', 'Credit Card',
        'CC Exp Date', 'CC Security Code', 'CC Provider', 'Email', 'Job',
        'IP Address', 'Language', 'Purchase Price'],
       dtype='object')
```

data[(data['CC Provider']=="Mastercard") & (data['Purchase Price']>50)]



	Address	Lot	AM or PM	Browser Info	Company	Credit Card	I Da
1	9374 Jasmine Spurs Suite 508\nSouth John, TN 8	28 rn	PM	Opera/8.93. (Windows 98; Win 9x 4.90; en-US) Pr	Fletcher, Richards and Whitaker	3337758169645356	11
18	461 Christopher Square\nWest Michaelchester, C	17 SB	PM	Mozilla/5.0 (X11; Linux i686; rv:1.9.6.20) Gec	Beard, Abbott and Pena	6011350184276270	12
31	USNS Alvarado\nFPO AA 27052-1231	26 Lh	PM	Opera/8.84. (X11; Linux i686; sl-Sl) Presto/2.9	Nicholson Group	4614997834548	03
35	93392 Webb Gardens Apt. 220\nLaurabury, AR 999	37 om	AM	Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_7	Mora Ltd	6011049630969815	09
90	431 Bowen Lights\nFergusonborough, MH 01362	31 tG	AM	Mozilla/5.0 (Windows NT 5.2; it-IT; rv:1.9.1.2	Copeland- Lee	639032576097	06
							-
9941	43757 Brown Lodge\nValerieberg, PR 88518	68 DG	AM	Mozilla/5.0 (compatible; MSIE 5.0; Windows NT	Jones, Williams and Dunn	6011508474487291	04
9948	95544 Johnson Isle Suite 939\nMichaelberg, RI	91 bW	AM	Opera/8.36. (X11; Linux x86_64; sI-SI) Presto/2	Fox- Peterson	4762924304307	03
9954	051 Samantha Wells Apt. 328\nScottborough, OR	11 vG	AM	Mozilla/5.0 (Macintosh; U; PPC Mac OS X 10_8_2	Mendoza, Zimmerman and Reilly	30232295374907	05
9981	087 Larson Roads Apt. 587\nRicestad, WI 98077	95 Kw	AM	Opera/9.63. (Windows NT 5.01; en-US) Presto/2.9	Ortiz, Schultz and Adams	4690863684605	01
9987	2754 Klein Mission\nJohnsonview, PA 64925-9624	65 Fg	РМ	Opera/9.26. (Windows CE; sl-SI) Presto/2.9.188	Curtis LLC	180027305702263	12