



E-COMMERCE PURCHASES INSIGHTS

Presented By : Vineet Patyal

TABLE OF CONTENTS

01

Business Overview

02

Data Analysis Tools

03

Business Queries &
Insights

04

Data Resource



| 01

⋮ BUSINESS OVERVIEW



Ecommerce Business Overview

Ecommerce business has revolutionized retail, enabling consumers to shop conveniently online. With the rise of digital platforms, purchasing goods and services has become seamless and accessible. Ecommerce purchases encompass a wide range of products, from electronics to groceries, offering consumers unparalleled convenience and choice. This shift in consumer behaviour has propelled the ecommerce sector's growth, driving increased sales and market expansion globally. As technology advances and consumer preferences evolve, ecommerce continues to thrive, reshaping the retail landscape and offering businesses new opportunities for growth and innovation in the digital marketplace.



! 02 DATA ANALYSIS TOOLS

TOOLS USED FOR ANALYSIS





03

BUSINESS QUERIES & INSIGHTS





BUSINESS QUERIES

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...)?
16. Which Credit Card Provider's cards has done the Most purchasing in terms of total value, Share top 5?

INSIGHTS

1. Display Top 10 Rows of The Dataset

INPUT : data.head(10)

OUTPUT

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
16629 Pace Camp Apt. 448\nAlexisborough, NE 77...	46 in	PM	Opera/9.56.(X11; Linux x86_64; sl-5) Presto/2...	Martinez-Hernan	6.01193E+15	Feb-20	900	JCB 16 digit	pduelap@yahoo.com	Scientist, product/process development	149.146.147.205	el	98.14
9374 Jasmine Spurs Suite 508\nSouth John, TN 8...	28 m	PM	Opera/8.93.(Windo ws 98; Win 9x 4.90; en-US) Pr...	Fletcher, Richards and Whitaker	3.33776E+15	Nov-18	561	Mastercard	anthony41@reed.com	Drilling engineer	15.160.41.51	fr	70.73
Unit 0065 Box 5052\nDPO AP 27450	94 vE	PM	Mozilla/5.0 (compatible; MSIE 9.0; Windows NT ...	Simpson, Williams and Pham	6.75958E+11	Aug-19	699	JCB 16 digit	amymiller@morales-harrison.com	Customer service manager	132.207.160.22	de	0.95
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	6.01158E+15	Feb-24	384	Discover	brent16@olson-robinson.info	Drilling engineer	30.250.74.19	es	78.04
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	6.01146E+15	Oct-25	678	Diners Club / Carte Blanche	christopherwright@g mail.com	Fine artist	24.140.33.94	es	77.82
7502 Powell Mission Apt. 768\nTravisland, VA 3...	21 XT	PM	Mozilla/5.0 (Macintosh; U; PPC Mac OS X 10_8_5...	Silva-Anderson	3.02462E+13	Jul-25	7169	Discover	ynghuyen@gmail.com	Fish farm manager	55.96.152.147	ru	25.15
93971 Conway Causeway\nAndersonburgh, AZ 75107	96 Xi	AM	Mozilla/5.0 (compatible; MSIE 7.0; Windows NT ...	Gibson and Sons	6.0114E+15	Jul-24	714	VISA 16 digit	olivia04@yahoo.com	Dancer	127.252.144.18	de	88.56
260 Rachel Plains Suite 366\nCastroberg, WV 24...	96 pG	PM	Linux i686) AppleWebKit/5350...	Marshall-Collins	5.61252E+11	Jun-25	256	VISA 13 digit	phillip48@parks.info	Event organiser	224.247.97.150	pt	44.25
2129 Dylan Burg\nNew Michelle, ME 28650	45 JN	PM	Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_7...	Galloway and Sons	1.80042E+14	Apr-24	899	JCB 16 digit	kdavis@rasmussen.co m	Financial manager	146.234.201.229	ru	59.54
3795 Dawson Extensions\nLake Tinalfort, ID 88739	15 Ug	AM	Mozilla/5.0 (X11; Linux i686; rv:1.9.7.20) Gec...	Rivera, Buchanan and Ramirez	4.39628E+12	Jan-17	931	American Express	gc Coleman@hunt-huerta.com	Forensic scientist	236.198.199.8	zh	95.63

2. Check Last 10 Rows of The Dataset

INPUT : data.tail(10)

OUTPUT

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
75731 Molly Springs\West Dannelle, VT 06034-5102	93 1y	PM	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_7_4...	Pace, Vazquez and Richards	8.69968E+14	Apr-24	877	JCB 15 digit	andersonmichael@sherman.biz	Early years teacher	54.170.3.185	ru	18.35
PSC 8165, Box B490\NPO AP 60327-0346	50 6A	AM	Mozilla/5.0 (compatible; MSIE 8.0; Windows NT ...	Snyder Inc	4.22158E+15	Feb-24	969	Voyager	kking@wise-ku.com	IT sales professional	254.25.31.156	el	25.93
885 Allen Mountains Apt. 230\Wallhaven, LA 16995	40 vH	PM	Mozilla/5.0 (Macintosh; PPC Mac OS X 10_6_5) A...	Wells Ltd	4.66483E+15	Oct-20	431	Discover	bberry@wright.net	Set designer	174.173.51.32	de	67.96
7555 Larson Locks Suite 229\Ellisburgh, MA 34...	72 J0	PM	Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_8...	Colon and Sons	3.00256E+13	Oct-25	629	Maestro	chelseawilliams@lopez.biz	Designer, exhibition/display	177.46.82.128	el	65.61
6276 Rojan Hollow\Lake Louis, WY 56410-7837	93 Iix	PM	Opera/9.68.(X11; Linux x86_64; x-SI) Presto/2...	Ritter-Smith	3.11219E+15	Jan-25	1823	Maestro	iroberts@gmail.com	Education officer, museum	242.44.112.18	zh	31.85
966 Castaneda Locks\West Juliafurt, CO 96415	92 X0	PM	Mozilla/5.0 (Windows NT 5.1) AppleWebKit/535 2...	Randall-Sloan	3.42945E+14	Mar-22	838	JCB 15 digit	iscott@wade-gamer.com	Printmaker	29.73.197.114	it	82.21
832 Curtis Dam Suite 785\North Edwardburgh, T...	41 Jf	AM	Mozilla/5.0 (compatible; MSIE 9.0; Windows NT ...	Hale, Collins and Wilson	2.10033E+14	Jul-25	207	JCB 16 digit	mary85@hotmail.com	Energy engineer	121.133.168.5 1	pt	25.63
Unit 4434 Box 6343\NDPO AE 28026-0283	74 Zh	AM	Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_7...	Anderson Ltd	6.01154E+15	May-21	1	VISA 16 digit	tyler16@gmail.com	Veterinary surgeon	156.210.0.254	el	83.98
0096 English Rest\Roystad, IA 12457	74 zL	PM	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_8_8...	Cook Inc	1.80003E+14	Nov-17	987	American Express	elizabethmoore@veid.net	Local government officer	55.78.26.143	es	38.84
40674 Barnett Stravenue\Grimmesville, WI 79682	64 Hr	AM	Mozilla/5.0 (X11; Linux i686; rv:1.9.5.20) Gec...	Greene Inc	4.13997E+15	Feb-19	302	JCB 15 digit	rachelford@vaughns.com	Embryologist clinical	176.119.198.1 99	el	67.59

3. Check Datatype of Each Column

INPUT : data.dtypes

OUTPUT

```
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

INPUT : `data.isnull().sum(axis=0)`

OUTPUT

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 in our Dataset?

INPUT

```
print('Number of rows', data.shape[0])  
print('Number of Columns', data.shape[1])
```

OUTPUT

```
Number of rows 10000  
Number of Columns 14
```

6. Highest and Lowest Purchase Prices.

INPUT `data['Purchase Price'].max()`

OUTPUT 99.99

INPUT `data['Purchase Price'].min()`

OUTPUT 0.0

7. Average Purchase Price

INPUT `data['Purchase Price'].mean()`

OUTPUT 50.347302

8. How many people have French 'fr' as their Language?

Using Length (Len) Method

INPUT

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

OUTPUT

1097

Using (count) Method

INPUT

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

OUTPUT

```
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

Using Length (Len) Method

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

OUTPUT 984

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

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

OUTPUT

```
2    amymiller@morales-harrison.com  
Name: Email, dtype: object
```


11. How many People have Mastercard as their Credit Card Provider and made a purchase above 50?

Using Length (Len) Method

INPUT

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

OUTPUT

405

Using (count) Method

INPUT

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

OUTPUT

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 the email of the person with the following Credit Card Number:
4664825258997302

INPUT `data[data['Credit Card']== 4664825258997302]['Email']`

OUTPUT

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

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

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

OUTPUT

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

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

Method 1 user def() function

INPUT

```
def fun():  
    count = 0  
    for date in data ['CC Exp Date']:  
        if date.split('/')[1]=='20':  
            count=count+1  
    print (count)
```

OUTPUT

```
fun()  
988
```

Method 2 (lambda x) function

INPUT

```
len(data[data['CC Exp Date'].apply(lambda x:x[3:] == '20')]) #we used x[3:] due to date value i.e. 02/20 (0 as 0 idx position,  
#2 as 1 position, / as 2 position, 2 as 3rd position  
#and : as all onwards 3rd position.
```

OUTPUT

```
988
```

15. What are the top 5 most popular email providers (e.g. gmail.com, yahoo.com, etc...)

Method 1 using Value counts function & Adding new column to data frame

INPUT

```
list1 = []  
for email in data['Email']:  
    list1.append(email.split('@')[1])  
  
data['temp'] = list1  
  
data['temp'].value_counts().head(5)
```

○
OUTPUT

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

15. What are the top 5 most popular email providers (e.g. gmail.com, yahoo.com, etc...)

Method 2 using apply (Lambda x) Function

INPUT

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

OUTPUT

```
data['temp'].value_counts().head(5)
```

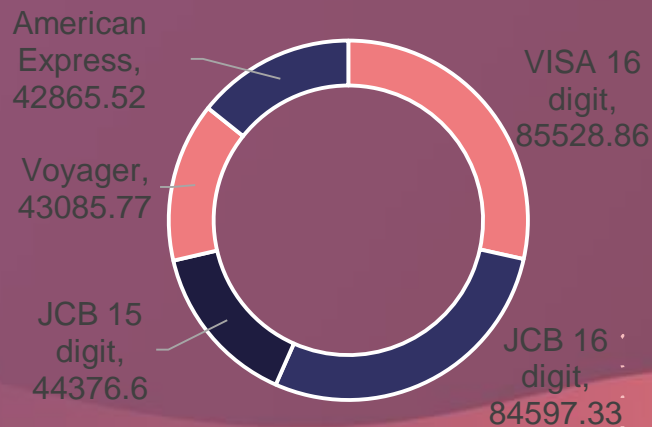
```
temp
hotmail.com      1638
yahoo.com        1616
gmail.com        1605
smith.com         42
williams.com      37
Name: count, dtype: int64
```

16. Which Credit Card Provider's cards has done the Most purchasing in terms of total value, Share top 5?

Using Group by & Sum function

INPUT `data.groupby('CC Provider')['Purchase Price'].sum().nlargest(5)`

OUTPUT



04|

DATA RESOURCE :

kaggle



THANK YOU!

