

[Solution] customer_purchase

September 3, 2022

The situation is, customers are providing some personal informations while purchasing stuff on-line or in-store. For some reasons, a client wants to know the answer to some of his questions from the dataset.

Task - 01 import desire library or libraries

```
[1]: # write your code here.  
import pandas as pd
```

Task - 02 load the data in a variable “customer”, data file name is ‘dataset.csv’.

```
[2]: customer = pd.read_csv('dataset.csv')
```

Task - 03 display first 5 rows of your data-set.

```
[3]: customer.head()
```

```
[3]:  prefix    first    last    email    gender    age  \  
0    Dr.      Ray    Morton    sebvajom@kol.km    Male    38  
1    Miss    Claudia    Rodriquez    acu@jatsot.ug    Female    51  
2    Miss    Harry    Meyer    zuz@lo.wf    Female    51  
3    Miss    Edith    Gilbert    hansohsi@jupec.md    Female    55  
4    Dr.      Lura    Murphy    webediti@je.st    Female    20  
  
      company    profession    phone  \  
0    Medtronic Inc.    Health Therapist    (987) 619-2695  
1    Ames Department Stores, Inc.    Health Therapist    (356) 736-7352  
2    CSX Corp.    Political Scientist    (539) 246-1806  
3    Murphy Oil Corporation    Transportation Manager    (984) 962-7494  
4    PETSMART Inc    Statistician    (902) 568-9748  
  
      postal province    cc_no    cc_exp    cc_type  \  
0    B6V 3W3    MB    5020000000000230    05/2018    Solo  
1    G7M 5F3    SK    5020000000000230    07/2028    Visa  
2    A0Z 6P9    NS    6300000000000000    02/2023    Switch  
3    P9I 9H3    YT    3530000000000000    02/2028    Maestro  
4    S1A 6K0    ON    4030000000000000    10/2025    Diners Club International
```

	price(CAD)	fav_color	ip	weekday	ampm	date
0	8.36	Blue	126.23.139.2	Sunday	pm	04/05/1930
1	68.31	Black	106.198.76.211	Tuesday	am	12/20/1926
2	34.65	Black	186.150.187.29	Wednesday	pm	08/20/1931
3	64.59	White	80.140.57.161	Saturday	am	06/18/2001
4	20.83	Yellow	211.103.43.41	Friday	pm	06/14/2045

Task - 04 How many entries your data have?

```
[4]: customer.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 30000 entries, 0 to 29999
Data columns (total 20 columns):
#   Column          Non-Null Count  Dtype
---  -
0   prefix          30000 non-null  object
1   first           30000 non-null  object
2   last            30000 non-null  object
3   email           30000 non-null  object
4   gender          30000 non-null  object
5   age             30000 non-null  int64
6   company         30000 non-null  object
7   profession      30000 non-null  object
8   phone           30000 non-null  object
9   postal          30000 non-null  object
10  province        30000 non-null  object
11  cc_no           30000 non-null  int64
12  cc_exp          30000 non-null  object
13  cc_type         30000 non-null  object
14  price(CAD)      30000 non-null  float64
15  fav_color       30000 non-null  object
16  ip              30000 non-null  object
17  weekday         30000 non-null  object
18  ampm            30000 non-null  object
19  date            30000 non-null  object
dtypes: float64(1), int64(2), object(17)
memory usage: 4.6+ MB
```

Task - 05 The max and min ages of your customer. Find mean of your customer.

```
[5]: print('Max. age of the customer is:', customer['age'].max())
      print('Min. age of the customer is:', customer['age'].min())
      print('Avg. age of the customer is:', round(customer['age'].mean(), 2))
```

Max. age of the customer is: 65

Min. age of the customer is: 18

Avg. age of the customer is: 41.55

Task - 06 the three most common customer's names

```
[6]: customer['first'].value_counts().head(3)
```

```
[6]: Willie      130
Francis      124
Eula         86
Name: first, dtype: int64
```

Task - 07 Two customers have the same phone number, find that phone number.

```
[7]: customer['phone'].value_counts().head().index[0]
```

```
[7]: '(263) 382-8004'
```

Task - 08 The customer(s) who has the same phone number

```
[8]: customer[customer['phone'] == '(263) 382-8004']
```

```
[8]:  prefix first  last          email gender age      company \
15  Mrs. Lilly Tyler  kofadu@itohi.tf  Female  38      CSX Corp.
16  Mrs. Peter  Cain  megkosig@anazeor.gn   Male  27  Campbell Soup Co.

      profession      phone  postal province      cc_no \
15  Structural Engineer  (263) 382-8004  V7K 1E3      ON  3000000000000000
16  Insurance Adjuster  (263) 382-8004  E8T 2B4      YT  3530000000000000

      cc_exp      cc_type  price(CAD)  fav_color      ip \
15  11/2022  Diners Club Carte Blanche      9.61  Yellow  74.124.37.227
16  03/2024      Solo      13.74    Black  25.207.141.135

      weekday ampm      date
15  Saturday   am  03/30/1985
16  Tuesday    am  08/05/1950
```

Task - 09 How many customers have profession “Structural Engineer”?

```
[9]: customer[customer['profession'] == 'Structural Engineer'].count()['profession']
```

```
[9]: 87
```

Task - 10 How many male customer are ‘Structural Engineer’?

```
[10]: customer[(customer['profession'] == 'Structural Engineer') &
↳ (customer['gender'] == 'Male')].count()['profession']
```

[10]: 43

Task - 11 the female Structural Engineer from province Alberta(AB)?

```
[11]: customer[(customer['profession'] == 'Structural Engineer') &
            ↪(customer['gender'] == 'Female') & (
                customer['province'] == 'AB')]
```

```
[11]:      prefix first      last      email gender age \
8858      Dr.   Roy   Stanley  apiunvo@ehasom.ir  Female  39
9058      Mr.   Lora  Kennedy   dennap@rabac.se  Female  51
24736  Mrs.   Nell  Richards    dob@me.ki  Female  18
29865  Mrs.   Don  McDaniel    naf@zudu.bj  Female  37

      company      profession      phone \
8858      Cincinnati Financial Corp.  Structural Engineer  (876) 758-2929
9058      BJ Services Company  Structural Engineer  (305) 786-6959
24736  Telephone & Data Systems Inc  Structural Engineer  (262) 681-5018
29865  Anadarko Petroleum Corporation  Structural Engineer  (238) 789-2825

      postal province      cc_no  cc_exp \
8858  N5A 6L2      AB  6300000000000000  02/2022
9058  N3R 4B7      AB  2010000000000000  06/2019
24736  N3Y 3G2      AB  3470000000000000  07/2028
29865  B4A 8Q9      AB  5610000000000000  12/2022

      cc_type  price(CAD) fav_color \
8858      InstaPayment      31.86      Red
9058  Diners Club United States & Canada      53.29      Black
24736      American Express      76.28      Blue
29865      Mastercard      62.71      Blue

      ip      weekday ampm      date
8858  165.249.159.57      Sunday      pm  12/25/1903
9058  42.216.243.206      Thursday      pm  06/06/2028
24736  91.16.30.156      Wednesday      pm  07/14/1951
29865  162.233.117.142      Sunday      pm  11/03/1989
```

Task - 12 The max, min and average spending

```
[12]: print('Max. spending:', customer['price(CAD)'].max())
      print('Min. spending:', customer['price(CAD)'].min())
      print('Avg. spending:', round(customer['price(CAD)'].mean(), 2))
```

```
Max. spending: 100.0
Min. spending: 0.0
Avg. spending: 49.99
```

Task - 13 Who did not spend anything? Company wants to send a deal to encourage the customer to buy stuff!

```
[13]: customer[customer['price(CAD)'] == 0.0]
```

```
[13]:      prefix first  last      email gender age      company \
5320    Dr.  Bruce  Bryan  ru@pubuspuh.cl  Male  59  Wal-Mart Stores Inc
10597  Mrs.  Flora  Clark    wad@me.com  Female  27    MetLife Inc.

      profession      phone  postal province      cc_no \
5320  Engineer Technician  (709) 446-8317  H6H 3Y0      NU  2010000000000000
10597    Cruise Director  (775) 373-6590  B2U 8K6      NB  3010000000000000

      cc_exp      cc_type  price(CAD) fav_color \
5320  04/2024      Bankcard      0.0      White
10597  05/2020  Diners Club Carte Blanche      0.0      Red

      ip  weekday ampm      date
5320  82.70.62.64  Wednesday  am  12/11/1900
10597  231.156.15.63  Thursday  pm  08/22/1904
```

Task - 14 As a loyalty reward, company wants to send thanks coupon to those who spent 100CAD or more, please find out the customers?

```
[14]: customer[customer['price(CAD)'] >= 100.0]
```

```
[14]:      prefix first  last      email gender age \
76    Mrs.  Gregory    Brown    hav@jek.gs  Female  31
21093  Mrs.    Cody  Christensen    get@jovu.ag  Male  28
24385  Miss  Lizzie    Dixon    goh@tuwjaz.gd  Female  38

      company      profession      phone \
76    E*Trade Group, Inc.    Novelist  (625) 537-8923
21093  National City Corp.  Hospital Administrator  (261) 737-3292
24385  FleetBoston Financial Co.  Compensation Analyst  (989) 239-1752

      postal province      cc_no  cc_exp  cc_type  price(CAD) \
76    X3S 4Q2      PE  601000000000000000  04/2018      Visa      100.0
21093  X0S 105      PE  201000000000000000  03/2024      JCB      100.0
24385  V8X 9V6      NB  630000000000000000  07/2019  Bankcard      100.0

      fav_color      ip  weekday ampm      date
76    Green  212.100.18.95  Monday  am  10/23/2063
21093  Green  168.48.19.165  Wednesday  pm  03/28/1955
24385  Blue  140.87.99.78  Saturday  pm  03/03/1983
```

Task - 15 How many emails are associated with this credit card number '5020000000000230'?

```
[15]: customer[customer['cc_no'] == 502000000000230]['email']
```

```
[15]: 0    sebvajom@kol.km
      1    acu@jatsot.ug
      Name: email, dtype: object
```

Task - 16 How many people use Visa as their Credit Card Provider?

```
[16]: sum(customer['cc_type'] == 'Visa')
```

```
[16]: 1721
```

Task - 17 The customer who spent 100 CAD using Visa?

```
[17]: customer[(customer['cc_type'] == 'Visa') & (customer['price(CAD)'] == 100.00)]
```

```
[17]:  prefix    first    last      email  gender  age      company \
76  Mrs.  Gregory  Brown  hav@jek.gs  Female   31  E*Trade Group, Inc.

      profession      phone  postal province      cc_no  cc_exp \
76  Novelist  (625) 537-8923  X3S 4Q2      PE  6010000000000000  04/2018

      cc_type  price(CAD) fav_color      ip weekday ampm      date
76  Visa      100.0      Green  212.100.18.95  Monday  am  10/23/2063
```

Task - 18 The two most common professions.

```
[18]: customer['profession'].value_counts().head(2)
```

```
[18]: Preschool Teacher      112
      Distribution Manager   107
      Name: profession, dtype: int64
```

Task - 19 Which day of the week, the store gets more customers?

```
[19]: customer['weekday'].value_counts().head(1)
```

```
[19]: Saturday      4376
      Name: weekday, dtype: int64
```

Task - 20 The most senior customer(s)?

```
[20]: customer[customer['age'] == customer['age'].max()]
```

```
[20]:  prefix    first    last      email  gender  age \
161  Miss  Minerva  Morales  anamimfo@socvobte.tz  Female   65
218  Mr.    Tyler    Rice    kumu@yahoo.com      Male   65
248  Mr.    Lou     Stone   mohzube@gmail.com    Male   65
```

365	Miss	Alberta	Nichols	naemozu@yahoo.com	Female	65
369	Mrs.	Johanna	Garcia	fugdihuw@yahoo.com	Male	65
...
29697	Dr.	Mabelle	Howell	wezru@hej.ac	Male	65
29722	Mrs.	Blake	Wheeler	jafmido@ko.mu	Male	65
29762	Miss	Beulah	Fleming	tijhohsuk@ririzen.cd	Male	65
29808	Dr.	Belle	Brock	otapihu@fofemi.pk	Female	65
29920	Miss	Roxie	Green	faniad@ud.ki	Male	65

		company	profession \
161	Anheuser-Busch Companies, Inc.		Importer
218	Kindred Healthcare Inc.		Computer Operator
248	Oracle Corp		Chemical Engineer
365	Kemet Corp.		Broadcaster
369	VF Corporation		Outplacement Specialist
...
29697	Cabot Corp		Education Adminator
29722	WestPoint Stevens Inc		Warehouse Manager
29762	Hughes Supply Inc.		Hospital Administrator
29808	M & T Bank Corporation		Bio-feedback Specialist
29920	Gap Inc.		Park Ranger

	phone	postal	province	cc_no	cc_exp \
161	(383) 327-5724	V7X 4A6	ON	2010000000000000	01/2028
218	(525) 280-8540	A6N 5C8	NS	5450000000000000	09/2020
248	(820) 471-5878	V3S 4B2	SK	5020000000000000	05/2024
365	(863) 432-8350	K2Y 5S2	NU	6210000000000000	02/2025
369	(622) 511-3495	L9U 8U8	NS	4030000000000000	02/2028
...
29697	(605) 289-3324	E6J 1P6	QC	6330000000000000	08/2018
29722	(661) 514-5657	V3G 7W3	MB	6370000000000000	07/2021
29762	(374) 515-8866	G8W 4S5	ON	4350000000000000	07/2024
29808	(936) 864-2101	H7U 6N3	YT	6330000000000000	02/2026
29920	(275) 641-2179	T0G 0V9	BC	6270000000000000	11/2024

	cc_type	price(CAD)	fav_color \
161	Diners Club Carte Blanche	47.50	Black
218	Switch	30.46	Blue
248	Solo	82.40	White
365	Bankcard	84.23	Black
369	InstaPayment	21.37	Blue
...
29697	JCB	80.39	Black
29722	Visa	52.01	White
29762	Diners Club United States & Canada	14.67	Blue
29808	Visa Electron	12.52	Red
29920	Discover Card	7.43	Red

	ip	weekday	ampm	date
161	2.133.75.235	Tuesday	am	12/05/1995
218	242.167.192.192	Friday	pm	08/24/2007
248	227.212.13.118	Wednesday	am	10/25/1992
365	24.39.70.9	Friday	am	06/05/2022
369	50.2.100.15	Tuesday	am	09/17/1952
...
29697	246.32.244.193	Monday	am	04/18/1932
29722	136.34.96.16	Thursday	am	07/28/2003
29762	85.114.159.98	Sunday	pm	05/01/2057
29808	160.33.185.137	Monday	am	02/20/1977
29920	134.150.236.40	Saturday	am	01/18/1967

[612 rows x 20 columns]

[]: