Gargi Sharma

12403355

Roll no:24

CA1

Data visualisation

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df1=pd.read csv("accounts.csv")
df2=pd.read csv("data_dictionary.csv")
df3=pd.read_csv("products.csv")
df4=pd.read csv("sales pipeline.csv")
df5=pd.read csv("sales teams.csv")
df3
          product series sales price
0
        GTX Basic
                                   550
                     GTX
1
          GTX Pro
                     GTX
                                  4821
2
       MG Special
                      MG
                                    55
3
      MG Advanced
                      MG
                                  3393
4
     GTX Plus Pro
                     GTX
                                  5482
5
  GTX Plus Basic
                     GTX
                                  1096
          GTK 500
                     GTK
                                 26768
df1
             account
                              sector year established
                                                        revenue
employees
    Acme Corporation
                           technolgy
                                                  1996 1100.04
2822
          Betasoloin
                             medical
                                                  1999
                                                         251.41
1
495
                             medical
                                                         647.18
            Betatech
                                                  1986
1185
3
          Bioholding
                             medical
                                                  2012
                                                         587.34
1356
             Bioplex
                             medical
                                                  1991
                                                         326.82
1016
. . .
          Zathunicon
                              retail
                                                  2010
                                                          71.12
80
```

```
144
81
      Zencorporation
                            technolgy
                                                     2011
                                                              40.79
142
82
               Zoomit entertainment
                                                     1992
                                                            324.19
978
83
              Zotware
                             software
                                                     1979
                                                           4478.47
13809
84
             Zumgoity
                              medical
                                                     1984
                                                             441.08
1210
   office location subsidiary of
     United States
0
                               NaN
1
     United States
                               NaN
2
              Kenya
                               NaN
3
        Philipines
                               NaN
4
     United States
                               NaN
                               . . .
80
     United States
                               NaN
81
              China
                               NaN
82
     United States
                               NaN
     United States
83
                               NaN
     United States
84
                               NaN
[85 rows x 7 columns]
df2
              Table
                                 Field \
0
          accounts
                               account
1
          accounts
                                sector
2
                     year_established
          accounts
3
           accounts
                               revenue
4
                             employees
           accounts
5
                      office location
          accounts
6
                         subsidiary of
          accounts
7
                               product
           products
8
           products
                                series
9
           products
                           sales_price
10
       sales teams
                           sales agent
11
       sales teams
                               manager
                       regional office
12
       sales teams
13
    sales pipeline
                       opportunity id
14
    sales pipeline
                           sales agent
15
    sales pipeline
                               product
16
    sales pipeline
                               account
                            deal_stage
17
    sales pipeline
18
    sales pipeline
                           engage date
19
    sales pipeline
                            close date
```

close value

20

sales pipeline

18 Date in v	peline st	tage (Pr e "Engag	Yenue (in Num Sugges Respecti U Tospecting ing" deal	Year Est millions ber of e Head Parent Producted reta Sal ve sales Regiona nique ic Sale Prod Comp > Engag stage v	employees Iquarters I company Iuct name It series I price I ces agent I office I entifier I sagent I uct name I any name I ing > I vas in	
			_			
opportur deal stage \	nity_id \	S	ales_agen	t	product	account
	C1I7A6R	M	loses Fras	e GTX F	Plus Basic	Cancity
1 Z0	9630YW0	Darce	el Schlech	t	GTXPro	Isdom
	C4QE1BX	Darce	l Schlech	t M	MG Special	Cancity
_	/1LWRNH	M	loses Fras	е	GTX Basic	Codehow
	E84CX40		Zane Lev	у	GTX Basic	Hatfan
Won						
 8795 9N	MIWFW5J	Versie	Hillebran	d MC	S Advanced	NaN
Prospecting						
Prospecting	SLKZ8FI		Hillebran		Advanced	NaN
8797 LI Prospecting	IB4KUZJ	Versie	Hillebran	d MC	6 Advanced	NaN
	BIUIUK0	Versie	Hillebran	d MC	Advanced	NaN
	I50NXJX	Versie	Hillebran	d MC	Advanced	NaN

	engage_date	close_date	close_value
0	2016-10-20	$2017 - \overline{0}3 - 01$	$\overline{1054.0}$
1	2016-10-25	2017-03-11	4514.0
2	2016-10-25	2017-03-07	50.0
3	2016-10-25	2017-03-09	588.0
4	2016-10-25	2017-03-02	517.0
8795	NaN	NaN	NaN
8796	NaN	NaN	NaN
8797	NaN	NaN	NaN
8798	NaN	NaN	NaN
8799	NaN	NaN	NaN

[8800 rows x 8 columns]

df5

0	sales_agent		regional_office Central
0	Anna Snelling	Dustin Brinkmann	
1	Cecily Lampkin	Dustin Brinkmann	Central
2	Versie Hillebrand	Dustin Brinkmann	Central
3	Lajuana Vencill	Dustin Brinkmann	Central
4	Moses Frase	Dustin Brinkmann	Central
5	Jonathan Berthelot	Melvin Marxen	Central
6	Marty Freudenburg	Melvin Marxen	Central
7	Gladys Colclough	Melvin Marxen	Central
8	Niesha Huffines	Melvin Marxen	Central
9	Darcel Schlecht	Melvin Marxen	Central
10	Mei-Mei Johns	Melvin Marxen	Central
11	Violet Mclelland	Cara Losch	East
12	Corliss Cosme	Cara Losch	East
13	Rosie Papadopoulos	Cara Losch	East
14	Garret Kinder	Cara Losch	East
15	Wilburn Farren	Cara Losch	East
16	Elizabeth Anderson	Cara Losch	East
17	Daniell Hammack	Rocco Neubert	East
18	Cassey Cress	Rocco Neubert	East
19	Donn Cantrell	Rocco Neubert	East
20	Reed Clapper	Rocco Neubert	East
21	Boris Faz	Rocco Neubert	East
22	Natalya Ivanova	Rocco Neubert	East
23	Vicki Laflamme	Celia Rouche	West
24	Rosalina Dieter	Celia Rouche	West
25	Hayden Neloms	Celia Rouche	West
26	Markita Hansen	Celia Rouche	West
27	Elease Gluck	Celia Rouche	West
28	Carol Thompson	Celia Rouche	West
29	James Ascencio	Summer Sewald	West
30	Kary Hendrixson	Summer Sewald	West
31	Kami Bicknell	Summer Sewald	West

32 33 34	Zane Levy Maureen Marcano Carl Lin	Summer	Sewalo Sewalo Sewalo	d	West West West		
df7=pedf8=pe	d.merge(df1,df4) d.merge(df6,df1) d.merge(df7,df3) .merge(df5,df8)						
0 1 2 3 4	sales_agent Anna Snelling Anna Snelling Anna Snelling Anna Snelling Anna Snelling	Dustin Bri Dustin Bri Dustin Bri Dustin Bri Dustin Bri	nkmann nkmann nkmann nkmann	regiona	l_office Central Central Central Central Central	account Betasoloin Betatech Betatech Betatech Betatech	\
6112 6113 6114 6115 6116	Maureen Marcano Maureen Marcano Maureen Marcano Maureen Marcano Maureen Marcano	Summer Summer Summer Summer	Sewald Sewald Sewald		West West West West West	Zoomit Zoomit Zoomit Zoomit Zoomit	
offic	sector yellocation \	ear_establi	shed i	revenue	employees	5	
0 State	_ medical		1999	251.41	495	5 United	
1 Kenya	medical		1986	647.18	1185	5	
2	medical		1986	647.18	1185	5	
Kenya 3	medical		1986	647.18	1185	5	
Kenya 4	medical		1986	647.18	1185	5	
Kenya 							
6112	entertainment		1992	324.19	978		
State			1992	324.19	978		
State	S						
6114 State	entertainment s		1992	324.19	978	3 United	
6115 State	entertainment s		1992	324.19	978	3 United	
6116 State	entertainment		1992	324.19	978	3 United	
	subsidiary_of opp e_date \	ortunity_id		produc	ct deal_	_stage	

NaM	\ \ \ \	D6020B	MG Adv	/anced	Prospecting	
ival	VV I	DOUZUN	no Au	ranceu	i i ospecting	
	N SK	Y95DVG	MG Sp	pecial	Lost	2016-
	V VT	5Y73GX	MG Sr	pecial	Won	2017-
			_			
	N HT	D2ME9C	GTX Plu	ıs Pro	Won	2017 -
Naf	N 8D	WOWB6Y	GTX	Basic	Lost	2017 -
• •	•					
	N 4B	PA17HF	GTX	Basic	Won	2017 -
	N WA	4TPQ5U	GTX Plus	Basic	Engaging	2017-
	u to	VODUER	CTV Dl.	ıc Dro	Lost	2017
	1 ID	Аорэоп	GIA PLL	IS PIU	LUST	2017-
	N RG	UC0V54	GTX Plus	Basic	Won	2017-
	N N8	SVXI48	GTX	Basic	Lost	2017-
2017-09-19 2017-11-14 rows x 18 co	- N 0 56 5370 0 628 N 0 882	aN I .0 I .0 G	MG MG TX TX TX TX TX TX	_price 3393 55 5482 550 550 1096 5482 1096 550		
sales_agent False False False False False False False False False	False	region	False False False False False False False False	Fals Fals Fals Fals Fals Fals	e False e False e False e False	
	Nai Nai Nai Nai Nai Nai Nai Nai Nai Nai	NaN SK NaN VT NaN HT NaN 8D NaN 4B NaN WA NaN NaN RG NaN NaN NaN 2017-03-22 0 0 2017-03-15 56 0 2017-06-12 5370 0 2017-07-3-16 0 0 2017-07-13 0 0 2017-09-19 882 2017-11-14 0 rows x 18 columns 1 0 0 rows x 18 columns 0 0 0 0 0 0	NaN SKY95DVG NaN VT5Y73GX NaN HTD2ME9C NaN 8DW0WB6Y NaN 4BPA17HF NaN WA4TPQ5U NaN IDA8D96H NaN RGUC0V54 NaN NaSVXI48 close_date close_value NaN NaN 2017-03-22 0.0 2017-04-12 5370.0 G 2017-06-12 5370.0 G 2017-03-16 0.0 G NaN NaN NaN 2017-05-28 628.0 G NaN NaN G 2017-07-13 0.0 G 2017-09-19 882.0 G 2017-11-14 0.0 G rows x 18 columns region False False False False False False False False False False	NaN SKY95DVG MG Sp NaN VT5Y73GX MG Sp NaN HTD2ME9C GTX Ptu NaN 8DW0WB6Y GTX NaN 4BPA17HF GTX NaN WA4TPQ5U GTX Ptus NaN RGUC0V54 GTX Ptus NaN NaSVXI48 GTX close_date close_value series sales_ NaN NaSVXI48 GTX close_date close_value series sales_ NaN NaSVXI48 GTX value series sales_ NaN NaN MG 2017-03-15 56.0 MG 2017-03-16 0.0 GTX 2017-05-28 628.0 GTX NaN NaN GTX 2017-07-13 0.0 GTX 2017-09-19 882.0 GTX 2017-11-14 0.0 GTX rows	NaN SKY95DVG MG Special NaN VT5Y73GX MG Special NaN HTD2ME9C GTX Plus Pro NaN 8DW0WB6Y GTX Basic NaN 4BPA17HF GTX Basic NaN WA4TPQ5U GTX Plus Basic NaN IDA8D96H GTX Plus Pro NaN RGUC0V54 GTX Plus Basic NaN N8SVXI48 GTX Basic close_date close_value series sales_price NaN N8SVXI48 GTX Basic close_date close_value series sales_price NaN RGUC0V54 GTX Plus Basic NaN NaSVXI48 GTX Basic 2017-03-22 0.0 MG 55 2017-03-15 56.0 MG 55 2017-03-16 0.0 GTX 5482 2017-03-16 0.0 GTX 550 NaN NaN GTX 1096 2017-07-13 0.0 GTX<	NaN SKY95DVG MG Special Lost NaN VT5Y73GX MG Special Won NaN HTD2ME9C GTX Plus Pro Won NaN 8DW0WB6Y GTX Basic Lost NaN 4BPA17HF GTX Plus Basic Won NaN WA4TPQ5U GTX Plus Basic Won NaN RGUC0V54 GTX Plus Basic Won NaN RGUC0V54 GTX Plus Basic Lost Close_date close_value series sales_price NaN NaN MG 55 2017-03-22 0.0 MG 55 2017-03-15 56.0 MG 55 2017-03-12 5370.0 GTX 5482 2017-03-13 0.0 GTX 550 NaN NaN GTX 550 NaN NaN GTX 1096 2017-07-13 0.0 GTX 5482 2017-09-

6115	False		.se	False		False	
6116	False	Fal	.se	False	e False	False	
cuboi	<pre>year_establi diary of \</pre>	shed	revenue	employees	office_loca	ation	
0		alse	False	False	F	alse	
True 1	F	alse	False	False	F	alse	
True 2	F	alse	False	False	F	alse	
True 3		alse	False	False		alse	
True							
4 True		alse	False	False	ŀ	alse	
6112	F	alse	False	False	F	alse	
True 6113 True	F	alse	False	False	F	alse	
6114	F	alse	False	False	F	alse	
True 6115	F	alse	False	False	F	alse	
True 6116	F	alse	False	False	F	alse	
True							
0 1 2 3 4	Fal Fal Fal	_id pr .se .se .se .se .se	roduct d False False False False False	eal_stage False False False False False	engage_date True False False False False	False False	\
6112 6113 6114 6115 6116	Fal Fal Fal Fal	.se .se .se .se .se	False False False False False	False False False False False	False False False False False	False False	
0 1 2 3 4 6112 6113	False False	Fals	se se se se	_price False False False False False False False False			

```
6114
             False
                     False
                                   False
6115
             False
                     False
                                   False
6116
            False
                     False
                                   False
[6117 rows x 18 columns]
df2
              Table
                                 Field \
0
          accounts
                               account
1
                                sector
          accounts
2
                     year_established
          accounts
3
          accounts
                               revenue
4
          accounts
                             employees
5
                      office_location
          accounts
6
                        subsidiary_of
          accounts
7
                               product
          products
8
          products
                                series
9
          products
                           sales price
10
       sales teams
                           sales agent
11
       sales teams
                               manager
12
       sales teams
                       regional office
13
                       opportunity id
    sales pipeline
14
    sales pipeline
                           sales agent
15
    sales pipeline
                               product
16
    sales pipeline
                               account
17
    sales pipeline
                            deal stage
18
    sales_pipeline
                           engage date
19
    sales pipeline
                            close date
20
    sales pipeline
                           close value
                                             Description
0
                                            Company name
1
                                                Industry
2
                                        Year Established
3
                   Annual revenue (in millions of USD)
4
                                    Number of employees
5
                                            Headquarters
6
                                          Parent company
7
                                            Product name
8
                                          Product series
9
                                 Suggested retail price
10
                                             Sales agent
11
                               Respective sales manager
12
                                         Regional office
13
                                       Unique identifier
14
                                            Sales agent
15
                                            Product name
16
                                            Company name
17
    Sales pipeline stage (Prospecting > Engaging >...
```

```
18
    Date in which the "Engaging" deal stage was in...
           Date in which the deal was "Won" or "Lost"
19
20
                                Revenue from the deal
df.dtypes
sales agent
                    object
manager
                    object
regional office
                    object
opportunity id
                    object
product
                    object
account
                    object
deal stage
                    object
engage date
                    object
close date
                    object
close value
                   float64
series
                    object
sales_price
                     int64
dtype: object
df.head()
                           manager regional office
     sales agent
                                                        account
sector \
O Anna Snelling Dustin Brinkmann
                                            Central
                                                     Betasoloin
medical
1 Anna Snelling
                  Dustin Brinkmann
                                            Central
                                                       Betatech
medical
2 Anna Snelling Dustin Brinkmann
                                            Central
                                                       Betatech
medical
3 Anna Snelling Dustin Brinkmann
                                            Central
                                                       Betatech
medical
4 Anna Snelling Dustin Brinkmann
                                            Central
                                                       Betatech
medical
                              employees office location subsidiary of
   year established
                     revenue
/
0
               1999
                      251.41
                                    495
                                           United States
                                                                   NaN
1
               1986
                      647.18
                                   1185
                                                                   NaN
                                                   Kenya
2
               1986
                      647.18
                                   1185
                                                                   NaN
                                                   Kenya
3
               1986
                      647.18
                                    1185
                                                   Kenya
                                                                   NaN
               1986
                      647.18
                                   1185
                                                   Kenya
                                                                   NaN
  opportunity id
                       product
                                 deal stage engage date
                                                          close date \
        WTD6020R
0
                   MG Advanced
                                Prospecting
                                                     NaN
                                                                 NaN
1
        SKY95DVG
                    MG Special
                                              2016-12-09
                                                          2017-03-22
                                       Lost
```

2 3 4		MG Special X Plus Pro GTX Basic	Won 20	17-02-19 2	017-03-15 017-06-12 017-03-16
clos	e value serie	s sales_price			
0	NaN M				
1	0.0 M				
2	56.0 M	G 55			
3	5370.0 GT				
4	0.0 GT	X 550			
df.tail	()				
coctor	sales_agen	t manage	r regional_o	office accou	nt
sector	∖ Haureen Marcan	o Summer Sewal	d	West Zoom	i +
enterta		o Sammer Seware	u	WC3C ZOOM	. .
	laureen Marcan	o Summer Sewal	d	West Zoom	it
enterta	inment				
	laureen Marcan	o Summer Sewal	d	West Zoom	it
enterta		- C	ـا	1-1-a-t 7-a-m	• 1
enterta	laureen Marcan	o Summer Sewal	u	West Zoom	11
	laureen Marcan	o Summer Sewal	h	West Zoom	i t
enterta			_		
_	ear_establish	ed revenue em	ployees offi	.ce_location	
subsidi 6112	ary_ot \ 19	92 324.19	978 Un	ited States	
NaN	19	92 324.19	970 011	illeu States	
6113	19	92 324.19	978 Un	ited States	
NaN	_				
6114	199	92 324.19	978 Un	ited States	
NaN	10.	00 004 10	070 11		
6115	199	92 324.19	978 Un	ited States	
NaN 6116	19	92 324.19	978 Un	ited States	
NaN	15	32 324113	370 011	itted States	
	portunity_id	product	deal_stage	engage_date	close_date
\ 6112	4BPA17HF	GTX Basic	Won	2017-05-14	2017-05-28
0112	4DFA17111	GIV Dasic	WOII	2017-03-14	2017-03-20
6113	WA4TPQ5U	GTX Plus Basic	Engaging	2017-06-13	NaN
6114	IDA8D96H	GTX Plus Pro	Lost	2017-07-05	2017-07-13
6115	RGUC0V54	GTX Plus Basic	Won	2017-07-11	2017-09-19
6116	N8SVXI48	GTX Basic	Lost	2017-09-23	2017-11-14

```
close value series
                           sales price
6112
            628.0
                     GTX
                                   550
6113
              NaN
                     GTX
                                  1096
              0.0
                                  5482
6114
                     GTX
            882.0
6115
                     GTX
                                  1096
6116
              0.0
                                   550
                     GTX
df['engage date'] =
pd.to datetime(df['engage date'].fillna(df['engage date'].mode()[0]))
df['close date'] =
pd.to datetime(df['close date'].fillna(df['close date'].mode()[0]))
df
          sales agent
                                 manager regional office
                                                              account \
0
        Anna Snelling
                       Dustin Brinkmann
                                                  Central
                                                           Betasoloin
1
        Anna Snelling Dustin Brinkmann
                                                  Central
                                                             Betatech
2
        Anna Snelling Dustin Brinkmann
                                                  Central
                                                             Betatech
3
        Anna Snelling
                       Dustin Brinkmann
                                                  Central
                                                             Betatech
4
        Anna Snelling
                       Dustin Brinkmann
                                                  Central
                                                             Betatech
     Maureen Marcano
                           Summer Sewald
                                                               Zoomit
6112
                                                     West
                           Summer Sewald
6113
     Maureen Marcano
                                                     West
                                                               Zoomit
6114
                           Summer Sewald
                                                               Zoomit
      Maureen Marcano
                                                     West
6115
                           Summer Sewald
                                                               Zoomit
      Maureen Marcano
                                                     West
6116 Maureen Marcano
                           Summer Sewald
                                                     West
                                                               Zoomit
             sector year_established revenue
                                                 employees
office location \
            medical
                                                        495
                                  1999
                                         251.41
                                                              United
0
States
            medical
                                  1986
                                         647.18
                                                       1185
Kenya
2
            medical
                                  1986
                                         647.18
                                                       1185
Kenya
3
            medical
                                  1986
                                         647.18
                                                       1185
Kenya
            medical
                                  1986
                                         647.18
                                                       1185
Kenya
                                                        . . .
6112 entertainment
                                  1992
                                         324.19
                                                        978
                                                              United
States
6113 entertainment
                                  1992
                                         324.19
                                                        978
                                                              United
States
6114 entertainment
                                  1992
                                         324.19
                                                        978
                                                              United
States
                                                        978
6115 entertainment
                                  1992
                                         324.19
                                                              United
States
```

```
1992
                                          324.19
                                                         978
                                                               United
6116 entertainment
States
     subsidiary of opportunity id
                                            product
                                                       deal stage
engage date \
                NaN
                          WTD6020R
                                        MG Advanced
                                                      Prospecting
                                                                    2017 -
07 - 19
1
                NaN
                          SKY95DVG
                                         MG Special
                                                             Lost 2016-
12-09
                                         MG Special
2
                NaN
                          VT5Y73GX
                                                              Won
                                                                    2017 -
01-09
                                       GTX Plus Pro
                                                              Won 2017-
3
                NaN
                          HTD2ME9C
02 - 19
                NaN
                          8DW0WB6Y
                                          GTX Basic
                                                             Lost 2017-
4
02-26
. . .
. . .
6112
                NaN
                          4BPA17HF
                                          GTX Basic
                                                              Won
                                                                    2017-
05 - 14
6113
                          WA4TP05U GTX Plus Basic
                                                         Engaging
                NaN
                                                                    2017 -
06 - 13
6114
                NaN
                          IDA8D96H
                                       GTX Plus Pro
                                                             Lost 2017-
07 - 05
6115
                NaN
                          RGUCOV54 GTX Plus Basic
                                                              Won 2017-
07-11
                          N8SVXI48
                                          GTX Basic
6116
                NaN
                                                             Lost 2017-
09 - 23
     close date
                 close value series
                                       sales_price
     2017-05-22
0
                          NaN
                                   MG
                                              3393
     2017-03-22
1
                          0.0
                                   MG
                                                 55
2
     2017-03-15
                         56.0
                                                 55
                                   MG
     2017-06-12
3
                       5370.0
                                  GTX
                                               5482
4
     2017-03-16
                          0.0
                                  GTX
                                                550
                                  . . .
                          . . .
6112 2017-05-28
                        628.0
                                               550
                                  GTX
6113 2017-05-22
                          NaN
                                  GTX
                                              1096
6114 2017-07-13
                          0.0
                                  GTX
                                              5482
6115 2017-09-19
                        882.0
                                              1096
                                  GTX
6116 2017-11-14
                                                550
                          0.0
                                  GTX
[6117 rows \times 18 columns]
df['engage date']=df['engage date'].fillna(df['engage date'].mode()
[0])
df['engage date']=pd.to datetime(df['engage date'])
df['close date']=df['close date'].fillna(df['close date'].mode()[0])
df['close date']=pd.to datetime(df['close date'])
df['year established']=pd.to datetime(df['year established'])
```

df						
0 1 2 3 4	sales_agen Anna Snellin Anna Snellin Anna Snellin Anna Snellin	g Dustin I g Dustin I g Dustin I g Dustin I	manager Brinkmann Brinkmann Brinkmann Brinkmann Brinkmann	Ce Ce Ce	ntral Be ntral ntral ntral	account \ tasoloin Betatech Betatech Betatech Betatech
6112 6113 6114 6115 6116	Maureen Marcan Maureen Marcan Maureen Marcan Maureen Marcan Maureen Marcan	o Summe o Summe o Summe o Summe	er Sewald er Sewald er Sewald er Sewald er Sewald		West West West West West	Zoomit Zoomit Zoomit Zoomit Zoomit
	sector		year_es	stablished	revenue	employees
0	medical	1970-01-01	00:00:00	.000001999	251.41	495
1	medical	1970-01-01	00:00:00	.000001986	647.18	1185
2	medical	1970-01-01	00:00:00	.000001986	647.18	1185
3	medical	1970-01-01	00:00:00	.000001986	647.18	1185
4	medical	1970-01-01	00:00:00	.000001986	647.18	1185
6112	entertainment	1970-01-01	00:00:00	.000001992	324.19	978
6113	entertainment	1970-01-01	00:00:00	.000001992	324.19	978
6114	entertainment	1970-01-01	00:00:00	.000001992	324.19	978
6115	entertainment	1970-01-01	00:00:00	.000001992	324.19	978
6116	entertainment	1970-01-01	00:00:00	.000001992	324.19	978
0 1 2 3 4 6112 6113 6114 6115	office_location United States Kenya Kenya Kenya Kenya United States United States United States		y_of oppor NaN NaN NaN NaN NaN NaN NaN NaN	rtunity_id WTD6020R SKY95DVG VT5Y73GX HTD2ME9C 8DW0WB6Y 4BPA17HF WA4TPQ5U IDA8D96H RGUC0V54	MG Ad MG S MG S GTX Pl GTX	Basic Basic Basic us Pro

```
6116
       United States
                                NaN
                                          N8SVXI48
                                                         GTX Basic
       deal stage engage date close date close value series
sales price
      Prospecting
                   2017-07-19 2017-05-22
                                                           MG
                                                   NaN
3393
                   2016-12-09 2017-03-22
                                                   0.0
                                                           MG
             Lost
1
55
2
              Won
                   2017-01-09 2017-03-15
                                                  56.0
                                                           MG
55
3
              Won
                   2017-02-19 2017-06-12
                                                5370.0
                                                          GTX
5482
4
             Lost 2017-02-26 2017-03-16
                                                   0.0
                                                          GTX
550
. . .
. . .
                   2017-05-14 2017-05-28
                                                 628.0
                                                          GTX
6112
              Won
550
         Engaging
                   2017-06-13 2017-05-22
                                                   NaN
                                                          GTX
6113
1096
6114
             Lost 2017-07-05 2017-07-13
                                                   0.0
                                                          GTX
5482
                   2017-07-11 2017-09-19
6115
              Won
                                                 882.0
                                                          GTX
1096
             Lost 2017-09-23 2017-11-14
6116
                                                   0.0
                                                          GTX
550
[6117 rows x 18 columns]
df['sales cycle duration'] = (df['close date'] -
df['engage date']).dt.days
plt.figure(figsize=(10, 6))
plt.hist(df['sales cycle duration'], bins=30, edgecolor='black')
plt.title('Distribution of Sales Cycle Durations')
plt.xlabel('Sales Cycle Duration (days)')
plt.ylabel('Frequency')
plt.show()
TypeError
                                           Traceback (most recent call
last)
File c:\Users\sharm\AppData\Local\Programs\Python\Python313\Lib\site-
packages\pandas\core\ops\array ops.py:218, in na arithmetic op(left,
right, op, is cmp)
    217 try:
            result = func(left, right)
--> 218
    219 except TypeError:
File c:\Users\sharm\AppData\Local\Programs\Python\Python313\Lib\site-
```

```
packages\pandas\core\computation\expressions.py:242, in evaluate(op,
a, b, use numexpr)
   240
           if use numexpr:
    241
                # error: "None" not callable
               return evaluate(op, op_str, a, b) # type:
--> 242
ignore[misc]
   243 return evaluate_standard(op, op_str, a, b)
File c:\Users\sharm\AppData\Local\Programs\Python\Python313\Lib\site-
packages\pandas\core\computation\expressions.py:73, in
evaluate standard(op, op str, a, b)
            store test result(False)
---> 73 return op(a, b)
TypeError: unsupported operand type(s) for -: 'str' and 'str'
During handling of the above exception, another exception occurred:
TypeError
                                          Traceback (most recent call
last)
Cell In[36], line 1
----> 1 df['sales cycle duration'] = (df['close date'] -
df['engage date']).dt.days
      2 plt.figure(figsize=(10, 6))
      3 plt.hist(df['sales cycle duration'], bins=30,
edgecolor='black')
File c:\Users\sharm\AppData\Local\Programs\Python\Python313\Lib\site-
packages\pandas\core\ops\common.py:76, in
unpack zerodim and defer.<locals>.new method(self, other)
                    return NotImplemented
     74 other = item from zerodim(other)
---> 76 return method(self, other)
File c:\Users\sharm\AppData\Local\Programs\Python\Python313\Lib\site-
packages\pandas\core\arraylike.py:194, in OpsMixin. sub (self,
other)
   192 @unpack zerodim and defer(" sub ")
   193 def sub (self, other):
--> 194
           return self. arith method(other, operator.sub)
File c:\Users\sharm\AppData\Local\Programs\Python\Python313\Lib\site-
packages\pandas\core\series.py:6135, in Series. arith method(self,
other, op)
   6133 def _arith_method(self, other, op):
           self, other = self. align for op(other)
   6134
           return base.IndexOpsMixin. arith method(self, other, op)
-> 6135
File c:\Users\sharm\AppData\Local\Programs\Python\Python313\Lib\site-
packages\pandas\core\base.py:1382, in
```

```
IndexOpsMixin. arith method(self, other, op)
            rvalues = np.arange(rvalues.start, rvalues.stop,
   1379
rvalues.step)
   1381 with np.errstate(all="ignore"):
            result = ops.arithmetic op(lvalues, rvalues, op)
   1384 return self. construct result(result, name=res name)
File c:\Users\sharm\AppData\Local\Programs\Python\Python313\Lib\site-
packages\pandas\core\ops\array_ops.py:283, in arithmetic_op(left,
right, op)
            bool arith check(op, left, right) # type: ignore[arg-
    279
type]
    281
            # error: Argument 1 to " na arithmetic op" has
incompatible type
            # "Union[ExtensionArray, ndarray[Any, Any]]"; expected
"ndarray[Any, Any]"
--> 283
           res values = na arithmetic op(left, right, op) # type:
ignore[arg-type]
    285 return res values
File c:\Users\sharm\AppData\Local\Programs\Python\Python313\Lib\site-
packages\pandas\core\ops\array ops.py:227, in na arithmetic op(left,
right, op, is cmp)
    219 except TypeError:
            if not is cmp and (
    221
                left.dtype == object or getattr(right, "dtype", None)
== obiect
    222
          ):
   (\ldots)
    225
                # Don't do this for comparisons, as that will handle
complex numbers
                # incorrectly, see GH#32047
    226
--> 227
                result = masked arith op(left, right, op)
    228
            else:
    229
                raise
File c:\Users\sharm\AppData\Local\Programs\Python\Python313\Lib\site-
packages\pandas\core\ops\array ops.py:163, in masked arith op(x, y,
op)
            # See GH#5284, GH#5035, GH#19448 for historical reference
    161
    162
            if mask.any():
                result[mask] = op(xrav[mask], yrav[mask])
--> 163
    165 else:
           if not is scalar(y):
TypeError: unsupported operand type(s) for -: 'str' and 'str'
df=pd.DataFrame(df[df["deal stage"]=="Won"].groupbv("sales agent")
["deal stage"].count()).reset index()
df.columns=["sales agent","Won"]
```

```
df.head()
plt.bar_label(plt.bar(df["sales_agent"],df["Won"]))
plt.xticks(rotation=90)
plt.show()
```

