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    GTX          550
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
6         GTK 500    GTK        26768

df1

            account       sector  year_established  revenue
employees  \
0   Acme Corporation    technolgy              1996  1100.04
2822
1         Betasoloin      medical              1999   251.41
495
2           Betatech      medical              1986   647.18
1185
3         Bioholding      medical              2012   587.34
1356
4            Bioplex      medical              1991   326.82
1016
..               ...          ...               ...      ...
...
80        Zathunicon       retail              2010    71.12
```

```
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
0     United States           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
83    United States           NaN
84    United States           NaN

[85 rows x 7 columns]

df2

             Table            Field  \
0          accounts          account
1          accounts           sector
2          accounts  year_established
3          accounts          revenue
4          accounts        employees
5          accounts  office_location
6          accounts     subsidiary_of
7          products          product
8          products           series
9          products      sales_price
10     sales_teams       sales_agent
11     sales_teams          manager
12     sales_teams  regional_office
13  sales_pipeline   opportunity_id
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...
19         Date in which the deal was "Won" or "Lost"
20                               Revenue from the deal
```

df4

```
      opportunity_id         sales_agent          product  account
deal_stage  \
0           1C1I7A6R         Moses Frase   GTX Plus Basic   Cancity
Won
1           Z063OYW0      Darcel Schlecht          GTXPro    Isdom
Won
2           EC4QE1BX      Darcel Schlecht      MG Special   Cancity
Won
3           MV1LWRNH         Moses Frase       GTX Basic   Codehow
Won
4           PE84CX40           Zane Levy       GTX Basic    Hatfan
Won
...              ...                 ...             ...      ...
...
8795        9MIWFW5J   Versie Hillebrand     MG Advanced       NaN
Prospecting
8796        6SLKZ8FI   Versie Hillebrand     MG Advanced       NaN
Prospecting
8797        LIB4KUZJ   Versie Hillebrand     MG Advanced       NaN
Prospecting
8798        18IUIUK0   Versie Hillebrand     MG Advanced       NaN
Prospecting
8799        8I5ONXJX   Versie Hillebrand     MG Advanced       NaN
Prospecting
```

```
       engage_date  close_date  close_value
0        2016-10-20  2017-03-01       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

            sales_agent           manager  regional_office
0         Anna Snelling  Dustin Brinkmann          Central
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           Zane Levy      Summer Sewald              West
33      Maureen Marcano     Summer Sewald              West
34           Carl Lin       Summer Sewald              West

df6=pd.merge(df1,df4)
df7=pd.merge(df6,df1)
df8=pd.merge(df7,df3)
df=pd.merge(df5,df8)
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
...	...	...	...	...
6112	Maureen Marcano	Summer Sewald	West	Zoomit
6113	Maureen Marcano	Summer Sewald	West	Zoomit
6114	Maureen Marcano	Summer Sewald	West	Zoomit
6115	Maureen Marcano	Summer Sewald	West	Zoomit
6116	Maureen Marcano	Summer Sewald	West	Zoomit

	sector	year_established	revenue	employees	office_location
0	medical	1999	251.41	495	United States
1	medical	1986	647.18	1185	Kenya
2	medical	1986	647.18	1185	Kenya
3	medical	1986	647.18	1185	Kenya
4	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
6115	entertainment	1992	324.19	978	United States
6116	entertainment	1992	324.19	978	United States

	subsidiary_of	opportunity_id	product	deal_stage	engage_date

```
0                NaN      WTD6O2OR       MG Advanced   Prospecting
NaN
1                NaN      SKY95DVG        MG Special          Lost  2016-
12-09
2                NaN      VT5Y73GX        MG Special           Won  2017-
01-09
3                NaN      HTD2ME9C      GTX Plus Pro           Won  2017-
02-19
4                NaN      8DWOWB6Y         GTX Basic          Lost  2017-
02-26
...              ...           ...               ...           ...
...
6112             NaN      4BPA17HF         GTX Basic           Won  2017-
05-14
6113             NaN      WA4TPQ5U  GTX Plus Basic      Engaging  2017-
06-13
6114             NaN      IDA8D96H      GTX Plus Pro          Lost  2017-
07-05
6115             NaN      RGUC0V54  GTX Plus Basic           Won  2017-
07-11
6116             NaN      N8SVXI48         GTX Basic          Lost  2017-
09-23

      close_date  close_value series  sales_price
0            NaN          NaN     MG         3393
1     2017-03-22          0.0     MG           55
2     2017-03-15         56.0     MG           55
3     2017-06-12       5370.0    GTX         5482
4     2017-03-16          0.0    GTX          550
...          ...          ...    ...          ...
6112  2017-05-28        628.0    GTX          550
6113         NaN          NaN    GTX         1096
6114  2017-07-13          0.0    GTX         5482
6115  2017-09-19        882.0    GTX         1096
6116  2017-11-14          0.0    GTX          550

[6117 rows x 18 columns]

df.isnull()

      sales_agent  manager  regional_office   account  sector  \
0           False    False            False     False   False
1           False    False            False     False   False
2           False    False            False     False   False
3           False    False            False     False   False
4           False    False            False     False   False
...           ...      ...              ...       ...     ...
6112        False    False            False     False   False
6113        False    False            False     False   False
6114        False    False            False     False   False
```

```
6115        False       False               False       False     False
6116        False       False               False       False     False

      year_established   revenue   employees   office_location
subsidiary_of  \
0                 False     False       False               False
True
1                 False     False       False               False
True
2                 False     False       False               False
True
3                 False     False       False               False
True
4                 False     False       False               False
True
...                 ...       ...         ...                 ...
...
6112              False     False       False               False
True
6113              False     False       False               False
True
6114              False     False       False               False
True
6115              False     False       False               False
True
6116              False     False       False               False
True

      opportunity_id   product   deal_stage   engage_date   close_date  \
0              False     False         False          True          True
1              False     False         False         False         False
2              False     False         False         False         False
3              False     False         False         False         False
4              False     False         False         False         False
...              ...       ...           ...           ...           ...
6112           False     False         False         False         False
6113           False     False         False         False          True
6114           False     False         False         False         False
6115           False     False         False         False         False
6116           False     False         False         False         False

      close_value   series   sales_price
0             True    False         False
1            False    False         False
2            False    False         False
3            False    False         False
4            False    False         False
...            ...      ...           ...
6112         False    False         False
6113          True    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          accounts             sector
2          accounts   year_established
3          accounts            revenue
4          accounts          employees
5          accounts    office_location
6          accounts       subsidiary_of
7          products            product
8          products             series
9          products        sales_price
10      sales_teams        sales_agent
11      sales_teams            manager
12      sales_teams    regional_office
13    sales_pipeline     opportunity_id
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...
19           Date in which the deal was "Won" or "Lost"
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()

```
      sales_agent            manager regional_office      account
sector  \
0  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

   year_established   revenue   employees office_location subsidiary_of
\
0              1999    251.41         495   United States           NaN

1              1986    647.18        1185           Kenya           NaN

2              1986    647.18        1185           Kenya           NaN

3              1986    647.18        1185           Kenya           NaN

4              1986    647.18        1185           Kenya           NaN


   opportunity_id       product    deal_stage engage_date  close_date  \
0        WTD6020R   MG Advanced   Prospecting         NaN         NaN
1        SKY95DVG    MG Special          Lost  2016-12-09  2017-03-22
```

```
2       VT5Y73GX     MG Special        Won  2017-01-09  2017-03-15
3       HTD2ME9C  GTX Plus Pro        Won  2017-02-19  2017-06-12
4       8DWOWB6Y     GTX Basic       Lost  2017-02-26  2017-03-16

   close_value series  sales_price
0         NaN     MG         3393
1         0.0     MG           55
2        56.0     MG           55
3      5370.0    GTX         5482
4         0.0    GTX          550
```

df.tail()

```
         sales_agent          manager regional_office account
sector  \
6112  Maureen Marcano  Summer Sewald            West   Zoomit
entertainment
6113  Maureen Marcano  Summer Sewald            West   Zoomit
entertainment
6114  Maureen Marcano  Summer Sewald            West   Zoomit
entertainment
6115  Maureen Marcano  Summer Sewald            West   Zoomit
entertainment
6116  Maureen Marcano  Summer Sewald            West   Zoomit
entertainment

      year_established  revenue  employees office_location
subsidiary_of  \
6112              1992   324.19        978   United States
NaN
6113              1992   324.19        978   United States
NaN
6114              1992   324.19        978   United States
NaN
6115              1992   324.19        978   United States
NaN
6116              1992   324.19        978   United States
NaN

     opportunity_id          product deal_stage engage_date  close_date
\
6112        4BPA17HF        GTX Basic        Won  2017-05-14  2017-05-28

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
6114          0.0    GTX         5482
6115        882.0    GTX         1096
6116          0.0    GTX          550

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
...                ...               ...             ...         ...
6112  Maureen Marcano     Summer Sewald            West      Zoomit
6113  Maureen Marcano     Summer Sewald            West      Zoomit
6114  Maureen Marcano     Summer Sewald            West      Zoomit
6115  Maureen Marcano     Summer Sewald            West      Zoomit
6116  Maureen Marcano     Summer Sewald            West      Zoomit

             sector  year_established  revenue  employees
office_location  \
0           medical              1999   251.41        495   United
States
1           medical              1986   647.18       1185
Kenya
2           medical              1986   647.18       1185
Kenya
3           medical              1986   647.18       1185
Kenya
4           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
6115  entertainment              1992   324.19        978   United
States
```

```
6116    entertainment              1992    324.19          978    United
States

     subsidiary_of opportunity_id          product   deal_stage
engage_date  \
0              NaN     WTD6O2OR     MG Advanced  Prospecting  2017-
07-19
1              NaN     SKY95DVG      MG Special         Lost  2016-
12-09
2              NaN     VT5Y73GX      MG Special          Won  2017-
01-09
3              NaN     HTD2ME9C    GTX Plus Pro          Won  2017-
02-19
4              NaN     8DWOWB6Y       GTX Basic         Lost  2017-
02-26
...            ...          ...             ...          ...
...
6112           NaN     4BPA17HF       GTX Basic          Won  2017-
05-14
6113           NaN     WA4TPQ5U  GTX Plus Basic     Engaging  2017-
06-13
6114           NaN     IDA8D96H    GTX Plus Pro         Lost  2017-
07-05
6115           NaN     RGUC0V54  GTX Plus Basic          Won  2017-
07-11
6116           NaN     N8SVXI48       GTX Basic         Lost  2017-
09-23

     close_date  close_value series  sales_price
0    2017-05-22          NaN     MG         3393
1    2017-03-22          0.0     MG           55
2    2017-03-15         56.0     MG           55
3    2017-06-12       5370.0    GTX         5482
4    2017-03-16          0.0    GTX          550
...         ...          ...    ...          ...
6112 2017-05-28        628.0    GTX          550
6113 2017-05-22          NaN    GTX         1096
6114 2017-07-13          0.0    GTX         5482
6115 2017-09-19        882.0    GTX         1096
6116 2017-11-14          0.0    GTX          550

[6117 rows x 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

            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
...                 ...                 ...              ...          ...
6112   Maureen Marcano      Summer Sewald              West       Zoomit
6113   Maureen Marcano      Summer Sewald              West       Zoomit
6114   Maureen Marcano      Summer Sewald              West       Zoomit
6115   Maureen Marcano      Summer Sewald              West       Zoomit
6116   Maureen Marcano      Summer Sewald              West       Zoomit

             sector              year_established  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


     office_location subsidiary_of opportunity_id         product  \
0      United States           NaN        WTD6O2OR      MG Advanced
1              Kenya           NaN        SKY95DVG       MG Special
2              Kenya           NaN        VT5Y73GX       MG Special
3              Kenya           NaN        HTD2ME9C     GTX Plus Pro
4              Kenya           NaN        8DWOWB6Y        GTX Basic
...              ...           ...             ...             ...
6112   United States           NaN        4BPA17HF        GTX Basic
6113   United States           NaN        WA4TPQ5U   GTX Plus Basic
6114   United States           NaN        IDA8D96H     GTX Plus Pro
6115   United States           NaN        RGUC0V54   GTX Plus Basic
```

```
6116     United States            NaN        N8SVXI48        GTX Basic

         deal_stage engage_date close_date   close_value series
sales_price
0      Prospecting  2017-07-19 2017-05-22           NaN     MG
3393
1              Lost  2016-12-09 2017-03-22           0.0     MG
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
...               ...         ...        ...           ...    ...
...
6112            Won  2017-05-14 2017-05-28         628.0    GTX
550
6113       Engaging  2017-06-13 2017-05-22           NaN    GTX
1096
6114           Lost  2017-07-05 2017-07-13           0.0    GTX
5482
6115            Won  2017-07-11 2017-09-19         882.0    GTX
1096
6116           Lost  2017-09-23 2017-11-14           0.0    GTX
550

[6117 rows x 18 columns]
```

```python
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:
--> 218     result = func(left, right)
    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
--> 242         return _evaluate(op, op_str, a, b)  # type:
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)
     72     _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)
     72             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):
   6134     self, other = self._align_for_op(other)
-> 6135     return base.IndexOpsMixin._arith_method(self, other, op)

File c:\Users\sharm\AppData\Local\Programs\Python\Python313\Lib\site-
packages\pandas\core\base.py:1382, in
```

```
IndexOpsMixin._arith_method(self, other, op)
   1379      rvalues = np.arange(rvalues.start, rvalues.stop,
rvalues.step)
   1381 with np.errstate(all="ignore"):
-> 1382      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)
   279      _bool_arith_check(op, left, right)  # type: ignore[arg-
type]
   281      # error: Argument 1 to "_na_arithmetic_op" has
incompatible type
   282      # "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:
   220      if not is_cmp and (
   221          left.dtype == object or getattr(right, "dtype", None)
== object
   222      ):
   (...)
   225          # Don't do this for comparisons, as that will handle
complex numbers
   226          #  incorrectly, see GH#32047
--> 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)
   161      # See GH#5284, GH#5035, GH#19448 for historical reference
   162      if mask.any():
--> 163          result[mask] = op(xrav[mask], yrav[mask])
   165 else:
   166      if not is_scalar(y):

TypeError: unsupported operand type(s) for -: 'str' and 'str'

df=pd.DataFrame(df[df["deal_stage"]=="Won"].groupby("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()
```