Mengukur Performa Penjualan Ritel Online Load Data

```
import pandas as pd
data = pd.read csv("C:\\Users\\mohra\\Downloads\\
online retail \overline{II}.csv")
data
        Invoice StockCode
                                                      Description
Quantity
         489434
                     85048
                            15CM CHRISTMAS GLASS BALL 20 LIGHTS
0
12
1
         489434
                    79323P
                                               PINK CHERRY LIGHTS
12
2
         489434
                    79323W
                                             WHITE CHERRY LIGHTS
12
                     22041
                                    RECORD FRAME 7" SINGLE SIZE
3
         489434
48
         489434
                     21232
                                  STRAWBERRY CERAMIC TRINKET BOX
24
                                    CHILDREN'S APRON DOLLY GIRL
1067366
         581587
                     22899
                     23254
                                   CHILDRENS CUTLERY DOLLY GIRL
1067367
         581587
1067368
                     23255
                                 CHILDRENS CUTLERY CIRCUS PARADE
         581587
1067369
         581587
                     22138
                                   BAKING SET 9 PIECE RETROSPOT
1067370
         581587
                      P<sub>0</sub>ST
                                                          POSTAGE
                  InvoiceDate
                                Price
                                       Customer ID
                                                             Country
         2009-12-01 07:45:00
0
                                 6.95
                                           13085.0
                                                     United Kingdom
1
         2009-12-01 07:45:00
                                           13085.0
                                                     United Kingdom
                                 6.75
2
         2009-12-01 07:45:00
                                 6.75
                                           13085.0
                                                     United Kingdom
3
                                           13085.0
         2009-12-01 07:45:00
                                 2.10
                                                     United Kingdom
         2009-12-01 07:45:00
                                 1.25
                                           13085.0
                                                     United Kingdom
1067366
         2011-12-09 12:50:00
                                 2.10
                                           12680.0
                                                              France
1067367
         2011-12-09 12:50:00
                                 4.15
                                           12680.0
                                                              France
1067368
         2011-12-09 12:50:00
                                 4.15
                                           12680.0
                                                              France
```

	2011-12-09 12:50:00	4.95	12680.0	France
	2011-12-09 12:50:00	18.00	12680.0	France
[1067371	rows x 8 columns]			

Section 1

Create New Feature: Year

<pre>data["InvoiceDate"] = pd.to_datetime(data["InvoiceDate"]) data["Year"] = data["InvoiceDate"].dt.year</pre>									
data									
		StockCode			Description				
Quantity 0	y \ 489434	85048	15CM CHR	ISTMAS GLASS	BALL 20 LIGHTS				
12									
1 12	489434	79323P		PINK	CHERRY LIGHTS				
2	489434	79323W		WHITE	CHERRY LIGHTS				
12			_						
3 48	489434	22041	R	RECORD FRAME 7	" SINGLE SIZE				
4	489434	21232	STR	AWBERRY CERAM	IC TRINKET BOX				
24									
						•			
1067366	581587	22899	C	HILDREN'S APR	ON DOLLY GIRL				
6 1067367	581587	23254	СН	ILDRENS CUTLE	RY DOLLY GIRL				
4 1067368	581587	23255	СПТІ	DDENC CUTLEDV	CIRCUS PARADE				
4	201201	23233	CHIL	DRENS CUILERT	CIRCUS PARADE				
1067369	581587	22138	BA	KING SET 9 PI	ECE RETROSPOT				
3 1067370	581587	POST			POSTAGE				
1	301307	1031			TOSTAGE				
		InvoiceDate	Price	Customer ID	Country	Year			
Θ	2009-12-	01 07:45:00	6.95	13085.0	United Kingdom	2009			
					_				
1	2009-12-	01 07:45:00	6.75	13085.0	United Kingdom	2009			
2	2009-12-	01 07:45:00	6.75	13085.0	United Kingdom	2009			
3	2009-12-	01 07:45:00	2.10	13085.0	United Kingdom	2009			
4	2009-12-	01 07:45:00	1.25	13085.0	United Kingdom	2009			

1067366 2011-12-09	12:50:00	2.10	12680.0	France	2011			
1067367 2011-12-09	12:50:00	4.15	12680.0	France	2011			
1067368 2011-12-09	12:50:00	4.15	12680.0	France	2011			
1067369 2011-12-09	12:50:00	4.95	12680.0	France	2011			
1067370 2011-12-09	12:50:00	18.00	12680.0	France	2011			
[1067371 rows x 9 columns]								

Filtering Data

```
sales = data[(data["Quantity"] > 0) &
(~data["Invoice"].str.startswith("C"))]
sales
        Invoice StockCode
                                                        Description
Quantity
          489434
                      85048
                              15CM CHRISTMAS GLASS BALL 20 LIGHTS
12
                                                 PINK CHERRY LIGHTS
1
          489434
                     79323P
12
                                               WHITE CHERRY LIGHTS
2
          489434
                     79323W
12
3
          489434
                      22041
                                      RECORD FRAME 7" SINGLE SIZE
48
4
          489434
                      21232
                                   STRAWBERRY CERAMIC TRINKET BOX
24
. . .
                                      CHILDREN'S APRON DOLLY GIRL
1067366
         581587
                      22899
                                    CHILDRENS CUTLERY DOLLY GIRL
1067367
         581587
                      23254
1067368
         581587
                      23255
                                  CHILDRENS CUTLERY CIRCUS PARADE
1067369
         581587
                      22138
                                    BAKING SET 9 PIECE RETROSPOT
                       P<sub>0</sub>ST
1067370
         581587
                                                             POSTAGE
1
                 InvoiceDate Price Customer ID
                                                              Country Year
        2009-12-01 07:45:00
0
                                 6.95
                                            13085.0 United Kingdom
                                                                        2009
```

_	2222 12 21					
1	2009-12-01	07:45:00	6.75	13085.0	United Kingdom	2009
2	2009-12-01	07:45:00	6.75	13085.0	United Kingdom	2009
3	2009-12-01	07:45:00	2.10	13085.0	United Kingdom	2009
4	2009-12-01	07:45:00	1.25	13085.0	United Kingdom	2009
1067366	2011-12-09	12:50:00	2.10	12680.0	France	2011
1067367	2011-12-09	12:50:00	4.15	12680.0	France	2011
1067368	2011-12-09	12:50:00	4.15	12680.0	France	2011
1067369	2011-12-09	12:50:00	4.95	12680.0	France	2011
1067370	2011-12-09	12:50:00	18.00	12680.0	France	2011
[1044420	orows x 9	columns]				

Create New Feature: Revenue

```
sales["Revenue"] = sales["Quantity"] * sales["Price"]
sales
C:\Users\mohra\AppData\Local\Temp\ipykernel 20548\605054434.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row indexer,col indexer] = value instead
See the caveats in the documentation:
https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#
returning-a-view-versus-a-copy
  sales["Revenue"] = sales["Quantity"] * sales["Price"]
        Invoice StockCode
                                                   Description
Quantity \
         489434
                    85048 15CM CHRISTMAS GLASS BALL 20 LIGHTS
12
1
         489434
                   79323P
                                            PINK CHERRY LIGHTS
12
                   79323W
                                           WHITE CHERRY LIGHTS
2
         489434
12
                                  RECORD FRAME 7" SINGLE SIZE
         489434
                    22041
48
                    21232
                                STRAWBERRY CERAMIC TRINKET BOX
         489434
```

24							
1067366 6	581587	22899	C	HILDREN'S APRO	ON DOLLY	/ GIRL	
1067367 4	581587	23254	CH	ILDRENS CUTLER	RY DOLLY	/ GIRL	
1067368 4	581587	23255	CHIL	DRENS CUTLERY	CIRCUS	PARADE	
1067369 3	581587	22138	BA	KING SET 9 PI	ECE RET	R0SP0T	
1067370 1	581587	POST			ſ	POSTAGE	
Revenue	Ir	nvoiceDate	Price	Customer ID		Country	Year
0 83.40	2009-12-01	07:45:00	6.95	13085.0	United	Kingdom	2009
1 81.00	2009-12-01	07:45:00	6.75	13085.0	United	Kingdom	2009
2 81.00	2009-12-01	07:45:00	6.75	13085.0	United	Kingdom	2009
3 100.80	2009-12-01	07:45:00	2.10	13085.0	United	Kingdom	2009
4 30.00	2009-12-01	07:45:00	1.25	13085.0	United	Kingdom	2009
1067366 12.60	2011-12-09	12:50:00	2.10	12680.0		France	2011
	2011-12-09	12:50:00	4.15	12680.0		France	2011
	2011-12-09	12:50:00	4.15	12680.0		France	2011
	2011-12-09	12:50:00	4.95	12680.0		France	2011
	2011-12-09	12:50:00	18.00	12680.0		France	2011
[1044420	o rows x 10	columns]					

Average of Revenue per Year

```
average_per_year = sales.groupby("Year")["Revenue"].mean()
average_per_year

Year
2009    18.713273
2010    19.908987
```

Interpretation

```
# Interpretation of Section 1
from IPython.display import Markdown

# Calculate the average revenue per year
average_per_year = sales.groupby("Year")["Revenue"].mean()

# Generate Markdown interpretation
interpretation = "### Interpretation of Section 1\n\n"
interpretation += "The table below shows the average annual revenue
for each year:\n\n"
interpretation += "| Year | Average Revenue |\n"
interpretation += "|-----|\n"

for year, average_revenue in average_revenue_per_year.items():
    interpretation += f"| {year} | {average_revenue:.2f} |\n"

# Display the interpretation in Markdown format
Markdown(interpretation)

<IPython.core.display.Markdown object>
```

Section 2

Filtering Data

Customers who finished their purchases

```
finished = sales.dropna(subset=["Customer ID"])
finished
        Invoice StockCode
                                                     Description
Quantity
         489434
                          15CM CHRISTMAS GLASS BALL 20 LIGHTS
0
                    85048
12
         489434
                   79323P
                                              PINK CHERRY LIGHTS
1
12
2
         489434
                   79323W
                                            WHITE CHERRY LIGHTS
12
                                   RECORD FRAME 7" SINGLE SIZE
3
         489434
                    22041
48
                                 STRAWBERRY CERAMIC TRINKET BOX
         489434
                    21232
4
24
```

1067366	581587	22899	C	HILDREN'S APRO	ON DOLLY	/ GIRL			
6	501507	22254	CUTI DDENC CUTI EDV DOLLY CID						
1067367	581587	23254	CHILDRENS CUTLERY DOLLY GIRL						
4 1067368	581587	23255	CHILDDENS CHILEDA CIDCUS DADADE						
4	301307	23233	CHILDRENS CUTLERY CIRCUS PARADE						
1067369	581587	22138	BAKING SET 9 PIECE RETROSPOT						
3	301307	22130	DANTING SET 9 PIECE KETKOSPOT						
1067370	581587	P0ST			F	POSTAGE			
1	00-00								
	In	voiceDate	Price	Customer ID		Country	Year		
Revenue									
0	2009-12-01	07:45:00	6.95	13085.0	United	Kingdom	2009		
83.40	2000 12 01	07 45 00		12005 0		17.	2000		
1	2009-12-01	07:45:00	6.75	13085.0	United	Kingdom	2009		
81.00 2	2009-12-01	07.45.00	6.75	13085.0	llni+od	Vinadom	2009		
81.00	2009-12-01	07:45:00	0.75	13003.0	onitted	Kingdom	2009		
3	2009-12-01	07.45.00	2.10	13085.0	United	Kingdom	2009		
100.80	2009-12-01	07.43.00	2.10	13003.0	OHITCEU	Kinguom	2009		
4	2009-12-01	07:45:00	1.25	13085.0	United	Kingdom	2009		
30.00									
	2011-12-09	12:50:00	2.10	12680.0		France	2011		
12.60						_			
	2011-12-09	12:50:00	4.15	12680.0		France	2011		
16.60	2011 12 00	12.50.00	4 1 5	12600 0		Гион оо	2011		
16.60	2011-12-09	12:50:00	4.15	12680.0		France	2011		
	2011-12-09	12.50.00	4.95	12680.0		France	2011		
14.85	2011-12-09	12.50.00	4.95	12000.0		Trance	2011		
	2011-12-09	12:50:00	18.00	12680.0		France	2011		
18.00		12.50.00	10100	1200010		Trance	2011		
[805620	rows x 10	columns]							

Customers who canceled their purchases

```
canceled = sales[sales['Invoice'].str.startswith('C')]
canceled

Empty DataFrame
Columns: [Invoice, StockCode, Description, Quantity, InvoiceDate,
Price, Customer ID, Country, Year, Revenue]
Index: []
```

Number of Finished and Canceled Transactions Each Year

```
finished_transaction = finished.groupby('Year').size()
canceled_transaction = canceled.groupby('Year').size()

finished_transaction

Year
2009     30761
2010     403094
2011     371765
dtype: int64

canceled_transaction

Series([], dtype: int64)
```

Cancellation Rate

```
canceled_rate = (canceled_transaction / (canceled_transaction +
finished_transaction) * 100)

canceled_rate

Year
2009 NaN
2010 NaN
2011 NaN
dtype: float64
```

Interpretation

```
# Interpretation of Section 2
from IPython.display import Markdown

# Customers who finished their purchases
finished = sales.dropna(subset=["Customer ID"])

# Customers who canceled their purchases
canceled = sales[sales['Invoice'].str.startswith('C')]

# Number of Finished and Canceled Transactions Each Year
finished_transaction = finished.groupby('Year').size()
canceled_transaction = canceled.groupby('Year').size()

# Cancellation Rate
canceled_rate = (canceled_transaction / (canceled_transaction +
finished_transaction) * 100)

# Generate Markdown interpretation
interpretation_section2 = "### Interpretation of Section 2\n\n"
```

```
interpretation section2 += "In Section 2, we analyze the transactions
of customers who either finished their purchases or canceled them.\n\
n"
interpretation section2 += f"There were {finished transaction.sum()}
unique customers who successfully finished their transactions, "
interpretation_section2 += f"and {canceled_transaction.sum()} unique
customers who canceled their purchases during the analyzed period.\n\
n"
interpretation section2 += "The total number of finished transactions
and canceled transactions each year are as follows:\n\n"
interpretation_section2 += "| Year | Finished Transactions | Canceled
Transactions |\n"
interpretation section2 +=
"|-----|\n"
for year in finished transaction.index:
    interpretation section2 += f" | {year} |
{finished transaction[year]} | {canceled transaction.get(year, 0)} |\
interpretation section2 += f"\nThe overall cancellation rate, which
represents the percentage of customers who canceled their "
interpretation section2 += f"orders out of all customers, is
{canceled rate.values[0]:.2f}%."
# Display the interpretation in Markdown format
Markdown(interpretation section2)
<IPython.core.display.Markdown object>
```