```
import pandas as pd
 In [4]:
 In [5]: df0 = pd.read_csv("./data/shipping_data_0.csv")
          df0.head()
 Out[5]:
             origin_warehouse destination_store product on_time product_quantity driver_identific
               d5566b15-b071-
                                 50d33715-4c77-
                                                                                       d8da0460-cf39
          0
                    4acf-8e8e-
                                     4dd9-8b9d-
                                                    lotion
                                                               True
                                                                                  59
                                                                                             4f38-9ff
                 c98433083b2d
                                    ff1ca372a2a2
                                                                                        6c9b4e344d8
                c42f0de8-b4f0-
                                  172eb8f3-1033-
                                                                                       293ccaec-6592
          1
                   4167-abd1-
                                      4fb6-b66b- windows
                                                               True
                                                                                  28
                                                                                            4f04-aae!
                 ae79e5e18eea
                                   d0df09df3161
                                                                                         3e238fe6261
               b145f396-de9b-
                                 65e4544d-42ae-
                                                                                       80988f09-91a3
          2
                    42f1-9cc9-
                                     4751-9580-
                                                      skis
                                                               True
                                                                                  63
                                                                                           4e1b-8e69
                 f5b52c3a941c
                                   bdcb90e5fcda
                                                                                         13551c53f31
                f4372224-759f-
                                 745bee4e-710c-
                                                                                        5f79b402-655
          3
                   43b3-bc83-
                                      4538-8df1-
                                                    bikes
                                                                                  47
                                                               True
                                                                                            4d8e-8ff3
                 ca6106bba1af
                                   5c146e1092a6
                                                                                         5ef05870e0a
               49d0edae-9091-
                                 425b7a1a-b744-
                                                                                       58beb5d3-98f8
          4
                   41bb-a08d-
                                      4c6b-898e-
                                                    candy
                                                              False
                                                                                  73
                                                                                           4077-a964
                 ab1c66bd08d5
                                   d424dd8cf18e
                                                                                         1f04f7cb11€
In [15]: df0.info()
        <class 'pandas.core.frame.DataFrame'>
        RangeIndex: 110 entries, 0 to 109
        Data columns (total 6 columns):
              Column
                                  Non-Null Count
                                                   Dtype
         0
              origin_warehouse
                                  110 non-null
                                                   object
         1
              destination_store 110 non-null
                                                   object
         2
              product
                                                   object
                                  110 non-null
         3
              on time
                                  110 non-null
                                                   bool
         4
              product_quantity
                                  110 non-null
                                                    int64
              driver_identifier 110 non-null
                                                   object
        dtypes: bool(1), int64(1), object(4)
        memory usage: 4.5+ KB
 In [7]: df1 = pd.read_csv("./data/shipping_data_1.csv")
          df1.head()
```

Out[7]:		shipment_identifier	product	on_time
	0	449263b4-6c93-4f19-8b6a-0d99a29fc637	pants	False
	1	449263b4-6c93-4f19-8b6a-0d99a29fc637	pants	False
	2	449263b4-6c93-4f19-8b6a-0d99a29fc637	pants	False
	3	449263b4-6c93-4f19-8b6a-0d99a29fc637	keyboards	False
	4	449263b4-6c93-4f19-8b6a-0d99a29fc637	keyboards	False

In [8]: df2 = pd.read_csv("./data/shipping_data_2.csv")
 df2.head()

Out[8]:		shipment_identifier	origin_warehouse	destination_store	driver_identifier
	0	449263b4-6c93-4f19- 8b6a-0d99a29fc637	bb75bf7d-c008-4267- bf92-6089cff5fe56	5e9405de-a078- 4b00-99c6- 96564568b63c	c12025e6-6f9c- 4728-8c3c- 9f840bde6f1a
	1	76e5b84a-9d09-4efb- 8b43-a0c932b958bb	372fd2b1-b2a7-4553- b6d7-426a1bc88e56	e34973c8-9ca9- 4a06-b497- 7a8b49625fc2	85b8d394-a67c- 48b6-b1de- 55be323ba622
	2	b541a47d-89b1-4805- 97d0-1988832321f1	469d957f-28ef-4eac- 956a-d2a42b06d3ab	fcadc756-61e9-41bb- 871b-d3546c5aa981	47bdfc40-f3db- 4678-b6a7- 43f1e1c2fd32
	3	3fc6b63d-27b4-408c- b3b3-e8e94a45b079	cd140190-a53b- 4660-a5b4- cc844a6506f0	89ba200c-ca90- 443a-b64f- 397bce091eae	5ae3e541-2098- 45b6-8d94- 35d176185606
	4	491ee4e8-be80-4f52- 802b-d8fe1a6bd487	c6addf8b-eea6-43b8- 9040-b5620b1a0d99	7aebe820-8478- 4a29-a606- 7c59af677e24	1f228b52-7165- 4d7f-a731- 3f7707aefb2a

In [9]: df_merged = pd.merge(df1,df2,how="inner",on="shipment_identifier")
 df_merged.head()

Out[9]:		shipment_identifier	product	on_time	origin_wareho	use dest	ination_store	driver_iden
	0	449263b4-6c93- 4f19-8b6a- 0d99a29fc637	pants	False	bb75bf7d-c0 4267-bf 6089cff5f	92-	9405de-a078- 4b00-99c6- 96564568b63c	c12025e6- 4728- 9f840bd
	1	449263b4-6c93- 4f19-8b6a- 0d99a29fc637	pants	False	bb75bf7d-c0 4267-bf 6089cff5f	92-	9405de-a078- 4b00-99c6- 96564568b63c	c12025e6- 4728- 9f840bd
	2	449263b4-6c93- 4f19-8b6a- 0d99a29fc637	pants	False	bb75bf7d-c0 4267-bf 6089cff5f	92-	9405de-a078- 4b00-99c6- 96564568b63c	c12025e6- 4728- 9f840bd
	3	449263b4-6c93- 4f19-8b6a- 0d99a29fc637	keyboards	False	bb75bf7d-c0 4267-bf 6089cff5f	92-	9405de-a078- 4b00-99c6- 96564568b63c	c12025e6- 4728- 9f840bd
	4	449263b4-6c93- 4f19-8b6a- 0d99a29fc637	keyboards	False	bb75bf7d-c0 4267-bf 6089cff5f	92-	9405de-a078- 4b00-99c6- 96564568b63c	c12025e6- 4728- 9f840bd
	4 (—
In [10]:		_final = df_merged[_final.head()	['origin_v	varehouse',	, 'dest	ination_	store', '	product',
Out[10]:		origin_warehouse	e dest	ination_stor	e product	on_time	driver_id	lentifier
	0	bb75bf7d-c008-4267- bf92-6089cff5fe56		e-a078-4b00 6564568b63	nants	False	c12025e6-6f9 8c3c-9f840	
	1	bb75bf7d-c008-4267- bf92-6089cff5fe56		e-a078-4b00 6564568b63	nants	False	c12025e6-6f9 8c3c-9f840	
	2	bb75bf7d-c008-4267- bf92-6089cff5fe56		e-a078-4b00 6564568b63	nants	False	c12025e6-6f9 8c3c-9f840	
	3	bb75bf7d-c008-4267- bf92-6089cff5fe56		e-a078-4b00 6564568b63	KEV/NOards	False	c12025e6-6f9 8c3c-9f840	
	4	bb75bf7d-c008-4267- bf92-6089cff5fe56		e-a078-4b00 6564568b63	keynoards	False	c12025e6-6f9 8c3c-9f840	
In [11]:	df_	_final.duplicated()	.sum()					
Out[11]:	np	.int64(66)						
In [12]:		_final_Product_quna _final_Product_quna	-		oupby([<mark>'origi</mark>	in_wareho	use', '	destinatio

<pre>df_f.head()</pre>			origin_warehouse	destination_store	product	on_time	driver_identifier	product_quantit
1 4daa-ae76- 4ba6-93ef- pencils True 4555-8434- a6b16580e873 20efa3c2-c498- 0167b0c3-60fd- 4908-8293- 4906-8293- 4908-8293- 6014b72de249 20efa3c2-c498- 0167b0c3-60fd- 4908-8293- 6014b72de249 20efa3c2-c498- 0167b0c3-60fd- 4908-8293- 6014b72de249 3 20efa3c2-c498- 4fbd-b378- toys False 45a23a59-09cb- 4906-8293- 6014b72de249 3 372fd2b1-b2a7- e34973c8-9ca9- baby food False 85b8d394-a67c- 426a1bc88e56 7a8b49625fc2 7a8b		0	4daa-ae76-	4ba6-93ef-	pants	True	4555-8434-	
2		1	4daa-ae76-	4ba6-93ef-	pencils	True	4555-8434-	
3		2	4908-8af4-	4fbd-b378-	eggs	False	4906-8293-	
4 4553-b6d7- 4a06-b497- food False 48b6-b1de-55be323ba622 **Total Control of State		3	4908-8af4-	4fbd-b378-	toys	False	4906-8293-	
Out[13]: origin_warehouse destination_store product on_time product_quantity driver_identifie 0 026ffd7c-cff2- 4daa-ae76- 6768d3283861 d33a7b5d-15e5- 4ba6-93ef- 4ba6-93ef- 4ba6-93ef- 6768d3283861 pants True 1 4555-8434 4b616580e87 1 026ffd7c-cff2- 4daa-ae76- 6768d3283861 d33a7b5d-15e5- 4ba6-93ef- 4ba6-93ef- 6768d3283861 pencils True 4 4555-8434 4555-8434 46b16580e87 2 20efa3c2-c498- 6768d3283861 0167b0c3-60fd- 4ba6-b378- 1ba0ceac39ff eggs False 3 4906-8293 6014b72de24 2 20efa3c2-c498- 181c76781912 0167b0c3-60fd- 1ba0ceac39ff toys False 3 4906-8293 6014b72de24 3 4908-8af4- 4fbd-b378- 181c76781912 45ba23a59-09cb 1ba0ceac39ff 45a23a59-09cb 1ba0ceac39ff 3 4906-8293 6014b72de24 3 372fd2b1-b2a7- 4 4553-b6d7- 436973c8-9ca9- 4a06-b497- baby 1baby 1cod False 1 48b6-b1de		4	4553-b6d7-	4a06-b497-	-	False	48b6-b1de-	
Out[13]: origin_warehouse destination_store product on_time product_quantity driver_identifie 0 026ffd7c-cff2- 4daa-ae76- 6768d3283861 d33a7b5d-15e5- 4ba6-93ef- 4ba6-93ef- 4ba6-93ef- 6768d3283861 pants True 1 4555-8434 4b616580e87 1 026ffd7c-cff2- 4daa-ae76- 6768d3283861 d33a7b5d-15e5- 4ba6-93ef- 4ba6-93ef- 6768d3283861 pencils True 4 4555-8434 4555-8434 46b16580e87 2 20efa3c2-c498- 6768d3283861 0167b0c3-60fd- 4ba6-b378- 1ba0ceac39ff eggs False 3 4906-8293 6014b72de24 2 20efa3c2-c498- 181c76781912 0167b0c3-60fd- 1ba0ceac39ff toys False 3 4906-8293 6014b72de24 3 4908-8af4- 4fbd-b378- 181c76781912 45ba23a59-09cb 1ba0ceac39ff 45a23a59-09cb 1ba0ceac39ff 3 4906-8293 6014b72de24 3 372fd2b1-b2a7- 4 4553-b6d7- 436973c8-9ca9- 4a06-b497- baby 1baby 1cod False 1 48b6-b1de		4 (•
026ffd7c-cff2- 0 d33a7b5d-15e5- 4ba6-93ef- 6768d3283861 pants True 1 4555-8434 a6b16580e87 1 026ffd7c-cff2- 6768d3283861 d33a7b5d-15e5- 4ba6-93ef- 4ba6-93ef- 6768d3283861 pencils True 4 4555-8434 a6b16580e87 2 026ffd7c-cff2- 4ba6-93ef- 6768d3283861 0167b0c3-60fd- 4ba6-93ef- 136bbcbf4946 pencils True 4 4555-8434 a6b16580e87 2 026fa3c2-c498- 4908-8af4- 4fbd-b378- f81c76781912 0167b0c3-60fd- fba0ceac39ff eggs False 3 4906-8293 6014b72de24 3 4908-8af4- 4fbd-b378- f81c76781912 4fbd-b378- fba0ceac39ff toys False 3 4906-8293 6014b72de24 3 4908-8af4- 4fbd-b378- f81c76781912 e34973c8-9ca9- fba0ceac39ff baby food False 85b8d394-a67c 1 4 4553-b6d7- 4553-b6d7- 4a06-b497- 4a06-b497- baby food False 1 48b6-b1de	In [13]:			duct_qunantity[['	origin_w	arehouse'	, 'destination_	_store', 'pr
0 4daa-ae76- 6768d3283861 4ba6-93ef- 136bbcbf4946 pants True 1 4555-8434 a6b16580e87 1 026ffd7c-cff2- 4daa-ae76- 6768d3283861 d33a7b5d-15e5- 4ba6-93ef- 136bbcbf4946 pencils True 4 4555-8434 4555-8434 2 20efa3c2-c498- 6768d3283861 0167b0c3-60fd- 14bd-b378- 158 eggs False 3 4906-8293 4906-8293 6014b72de24 2 20efa3c2-c498- 181c76781912 0167b0c3-60fd- 15b0ceac39ff toys False 3 4906-8293 4906-8293 6014b72de24 3 4908-8af4- 181c76781912 4fbd-b378- 15b0ceac39ff toys False 3 4906-8293 6014b72de24 3 372fd2b1-b2a7- 4353-b6d7- e34973c8-9ca9- 4a06-b497- 4a06-b497- baby food False 1 48b6-b1de	Out[13]:		origin warehouse	doctination stars	_			
1 4daa-ae76-6768d3283861 4ba6-93ef- pencils True 4 4555-8434 a6b16580e87 2 20efa3c2-c498- 4908-8af4- 4fbd-b378- f81c76781912 0167b0c3-60fd- fba0ceac39ff eggs False 3 4906-8293 6014b72de24 3 4908-8af4- 4fbd-b378- f81c76781912 toys False 3 4906-8293 6014b72de24 4 372fd2b1-b2a7- 4553-b6d7- 4a06-b497- food baby food False 85b8d394-a67c 4 4553-b6d7- 4a06-b497- food False 1 48b6-b1de				destination_store	product	on_time	product_quantity	driver_identifie
2 4908-8af4- 4fbd-b378- eggs False 3 4906-8293 f81c76781912 fba0ceac39ff 6014b72de24 20efa3c2-c498- 0167b0c3-60fd- 45a23a59-09cb 3 4908-8af4- 4fbd-b378- toys False 3 4906-8293 f81c76781912 fba0ceac39ff 6014b72de24 372fd2b1-b2a7- e34973c8-9ca9- baby food False 1 48b6-b1de		0	026ffd7c-cff2- 4daa-ae76-	d33a7b5d-15e5- 4ba6-93ef-	<u> </u>		· · · ·	6e905a32-46a8 4555-8434
3 4908-8af4- 4fbd-b378- toys False 3 4906-8293 f81c76781912 fba0ceac39ff 6014b72de24 372fd2b1-b2a7- e34973c8-9ca9- baby food False 1 48b6-b1de			026ffd7c-cff2- 4daa-ae76- 6768d3283861 026ffd7c-cff2- 4daa-ae76-	d33a7b5d-15e5- 4ba6-93ef- 136bbcbf4946 d33a7b5d-15e5- 4ba6-93ef-	pants	True	1	6e905a32-46a8 4555-8434 a6b16580e87 6e905a32-46a8 4555-8434
4 4553-b6d7- 4a06-b497- baby food False 1 48b6-b1de		1	026ffd7c-cff2- 4daa-ae76- 6768d3283861 026ffd7c-cff2- 4daa-ae76- 6768d3283861 20efa3c2-c498- 4908-8af4-	d33a7b5d-15e5- 4ba6-93ef- 136bbcbf4946 d33a7b5d-15e5- 4ba6-93ef- 136bbcbf4946 0167b0c3-60fd- 4fbd-b378-	pants	True	1	6e905a32-46a8 4555-8434 a6b16580e87 6e905a32-46a8 4555-8434 a6b16580e87 45a23a59-09cb 4906-8293
		1 2	026ffd7c-cff2- 4daa-ae76- 6768d3283861 026ffd7c-cff2- 4daa-ae76- 6768d3283861 20efa3c2-c498- 4908-8af4- f81c76781912 20efa3c2-c498- 4908-8af4-	d33a7b5d-15e5- 4ba6-93ef- 136bbcbf4946 d33a7b5d-15e5- 4ba6-93ef- 136bbcbf4946 0167b0c3-60fd- 4fbd-b378- fba0ceac39ff 0167b0c3-60fd- 4fbd-b378-	pants pencils eggs	True True False	1	6e905a32-46a8 4555-8434 a6b16580e87 6e905a32-46a8 4555-8434 a6b16580e87 45a23a59-09cb 4906-8293 6014b72de24 45a23a59-09cb 4906-8293
1		1 2 3	026ffd7c-cff2- 4daa-ae76- 6768d3283861 026ffd7c-cff2- 4daa-ae76- 6768d3283861 20efa3c2-c498- 4908-8af4- f81c76781912 20efa3c2-c498- 4908-8af4- f81c76781912 372fd2b1-b2a7- 4553-b6d7-	d33a7b5d-15e5- 4ba6-93ef- 136bbcbf4946 d33a7b5d-15e5- 4ba6-93ef- 136bbcbf4946 0167b0c3-60fd- 4fbd-b378- fba0ceac39ff 0167b0c3-60fd- 4fbd-b378- fba0ceac39ff e34973c8-9ca9- 4a06-b497-	pants pencils eggs toys baby	True False False	1 4 3	6e905a32-46a8 4555-8434 a6b16580e87 6e905a32-46a8 4555-8434 a6b16580e87 45a23a59-09ck 4906-8293 6014b72de24 45a23a59-09ck 4906-8293 6014b72de24 85b8d394-a67c 48b6-b1de
<pre>In [14]: df_f.info()</pre>		1 2 3	026ffd7c-cff2- 4daa-ae76- 6768d3283861 026ffd7c-cff2- 4daa-ae76- 6768d3283861 20efa3c2-c498- 4908-8af4- f81c76781912 20efa3c2-c498- 4908-8af4- f81c76781912 372fd2b1-b2a7- 4553-b6d7-	d33a7b5d-15e5- 4ba6-93ef- 136bbcbf4946 d33a7b5d-15e5- 4ba6-93ef- 136bbcbf4946 0167b0c3-60fd- 4fbd-b378- fba0ceac39ff 0167b0c3-60fd- 4fbd-b378- fba0ceac39ff e34973c8-9ca9- 4a06-b497-	pants pencils eggs toys baby	True False False	1 4 3	6e905a32-46a8 4555-8434 a6b16580e87 6e905a32-46a8 4555-8434 a6b16580e87 45a23a59-09ck 4906-8293 6014b72de24 45a23a59-09ck 4906-8293 6014b72de24 85b8d394-a67c 48b6-b1de

file:///D:/NEU/Forage/Walmart Global Tech- advance software engineering/Task 4/forage-walmart-task-4/Prepare Data.html

```
<class 'pandas.core.frame.DataFrame'>
       RangeIndex: 44 entries, 0 to 43
       Data columns (total 6 columns):
            Column
                              Non-Null Count Dtype
           -----
                               -----
                                             ----
        0
            origin_warehouse 44 non-null
                                              object
        1
            destination_store 44 non-null
                                              object
        2
            product
                              44 non-null
                                              object
        3
                                              bool
            on time
                              44 non-null
        4
            product_quantity 44 non-null
                                              int64
            driver_identifier 44 non-null
                                              object
       dtypes: bool(1), int64(1), object(4)
       memory usage: 1.9+ KB
In [43]: df_f.to_csv('./data/shipping_data_1&2_formated.csv', index=False)
```