Executing 1 NF in Barens Table

Barens table has multivalued column in prod_name

Hence we have separated comma values to a new column called volume.

```
ba = barens.drop (['prod_name'], axis = 1)

df['prod_volume'] = df['prod_name'].str.split(',')

df

df_models = pd.DataFrame(df['prod_volume'].tolist()).fillna('').add_prefix('column')

df_models.drop(['column2'], axis=1)
```

	column0	column1
0	Heartstopper	Volume 2
1	Heartstopper	Volume 1

```
In [106]: dffinal = pd.concat([ba, df_models], axis=1)

In [107]: dffinal
```

Out[107]:

	category	Price	webiste	column0	column1	column2
0	book	12.99	barnes	Heartstopper	Volume 2	
1	book	12.99	barnes	Heartstopper	Volume 1	
2	book	19.99	barnes	The Mothers and Daughters of the Bible Speak: ...		
3	book	14.99	barnes	Book Lovers		
4	book	14.99	barnes	Knight Owl		
5	book	31.99	barnes	Fifty Shades Freed 10th Anniversary Edition		
6	book	24.99	barnes	This Will Not Pass: Trump	Biden	and the Battle for America's Future
7	book	22.99	barnes	Finding Me		
8	book	5.00	barnes	Happy Dreamer		
9	book	13.99	barnes	It Starts with Us: A Novel		
10	book	12.99	barnes	The Seven Husbands of Evelyn Hugo: A Novel		
11	book	12.99	barnes	It Ends with Us		
12	book	19.99	barnes	22 Seconds		
13	book	20.99	barnes	Killing the Killers: The Secret War Against Te...		
14	book	16.99	barnes	Lightlark		

```
In [110]: dffinal.to_sql(name = 'barenfinal', con = engine, if_exists='append', index = False)
Out[110]: 21
```

```
In [93]: dffinal1.to_sql(name = 'barnesfinal', con = engine, if_exists='append', index = False)
Out[93]: 21
```
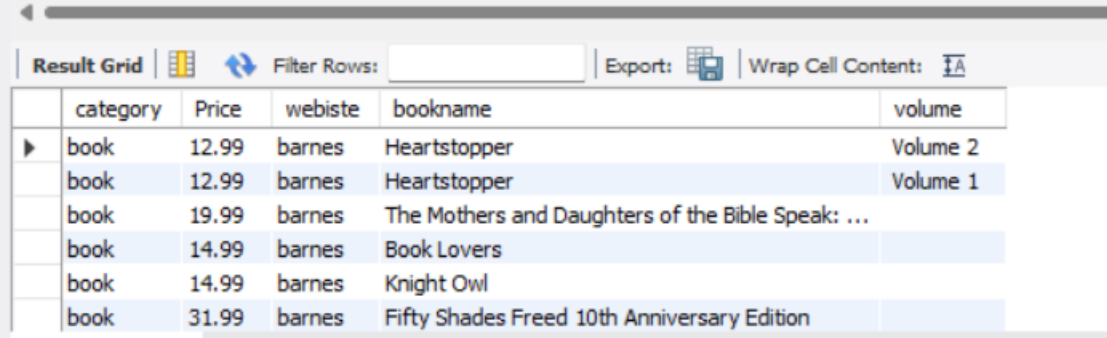
```
In [82]: df_final = df['prod_name'],df['Price']
```

```
In [83]: df_final
Out[83]: (0                         Heartstopper, Volume 2
         1                         Heartstopper, Volume 1
```

After inserting the data in table we need to rename the columns and hence we can see the result in this format .

```
118
119 ●   select * from barenfinal;
```

category	Price	webiste	bookname	volume
book	12.99	barnes	Heartstopper	Volume 2
book	12.99	barnes	Heartstopper	Volume 1
book	19.99	barnes	The Mothers and Daughters of the Bible Speak: ...	
book	14.99	barnes	Book Lovers	
book	14.99	barnes	Knight Owl	
book	31.99	barnes	Fifty Shades Freed 10th Anniversary Edition	

barenfinal 25 ✕

Output

Here we can we that the volume and book name are separated.

2 NF

As category,prod_specifcation,url and is_amazon_seller is independent of Uniq_id and it is dependent on prod_name we are normalizing and decomposing it to another table.

```
n [116]: amazon_df.drop(['Uniq_Id','Brand_Name','Model_Number','Shipping_Weight','Selling_Price'], axis=1)
```

out[116]:

	Product_Name	Category	Product_Specification	Product_Url	Is_Amazon_Seller
0	DB Longboards CoreFlex Crossbow 41" Bamboo Fib...	Sports & Outdoors | Outdoor Recreation | Skate...	Shipping Weight: 10.7 pounds (View shipping ra...	https://www.amazon.com/DB-Longboards-CoreFlex-...	Y
1	Electronic Snap Circuits Mini Kits Classpack, ...	Toys & Games | Learning & Education | Science ...	Product Dimensions: 14.7 x 11.1 x 10.2...	https://www.amazon.com/Electronic-Circuits-Cla...	Y
2	3Doodler Create Flexy 3D Printing Filament Ref...	Toys & Games | Arts & Crafts | Craft Kits	ProductDimensions:10.3x3.4x0.8inches|ItemWeigh...	https://www.amazon.com/3Doodler-Plastic-Innova...	Y
3	Guillow Airplane Design Studio with Travel Cas...	Toys & Games | Hobbies | Models & Model Kits |...	ProductDimensions:3.5x6.2x13inches|ItemWeight:...	https://www.amazon.com/Guillow-Airplane-Design...	Y
4	Woodstock- Collage 500 pc Puzzle	Toys & Games | Puzzles | Jigsaw Puzzles	ProductDimensions:1.9x8x10inches|ItemWeight:13...	https://www.amazon.com/Woodstock-Collage-500-p...	Y
...	...	...	...	...	...
9997	Remedia Publications REM536B Money Activity Bo...	Toys & Games | Learning & Education | Counting...	Product Dimensions: 11 x 8.5 x 0.4 inc...	https://www.amazon.com/Remedia-Publications-RE...	Y
9998	Trends International NFL La Chargers HG - Mobi...	Toys & Games | Arts & Crafts	ProductDimensions:3x6.5x0.3inches|ItemWeight:0...	https://www.amazon.com/Trends-International-NF...	Y
9999	NewPath Learning 10 Piece Science Owls and Owl...	Office Products | Office & School Supplies | E...	Product Dimensions: 18 x 11.9 x 0.7 in...	https://www.amazon.com/NewPath-Learning-Scienc...	Y
10000	Disney Princess Do It Yourself Braid Set	Toys & Games | Arts & Crafts | Craft Kits	ProductDimensions:7x7x2inches|ItemWeight:6.1ou...	https://www.amazon.com/Disney-Princess-Yoursel...	Y

```
n [136]: walmart_p =  walmart_df.drop(['Uniq_Id','Crawl_Timestamp','Sale_Price','Item_Number','Gtin','webiste'],axis = 1)
```

```
n [138]: walmart_p
```

ut[138]:

	Product_Url	Product_Name	Brand	Product_inovation	Category	Availa
0	https://www.walmart.com/ip/La-Costena-Chipotle...	La Costena Chipotle Peppers, 7 OZ (Pack of 12)	La CosteÃ¯Â¿Â½Ã¯Â¿Â½a	NaN	Food | Meal Solutions, Grains & Pasta | Canned...	T
1	https://www.walmart.com/ip/Equate-Triamcinolon...	Equate Triamcinolone Acetonide Nasal Allergy S	Equate	NaN	Health | Equate | Equate Allergy | Equate Sinu	T

For apple table we are filtering out columns which are partially or fully dependent on model name.

```
apple_pro =  apple_pro_df.drop(['pid','country_code','scraped_date'],axis = 1)
```

```
apple_pro
```

	model	price	country	region	incomegroup
0	iPhone 13	From A$1,199	Australia	East Asia & Pacific	High income
1	iPhone SE	From A$679	Australia	East Asia & Pacific	High income
2	iPhone 12	From A$999	Australia	East Asia & Pacific	High income
3	AirPods(3rd generation)	A$279	Australia	East Asia & Pacific	High income
4	AirPods Pro	A$399	Australia	East Asia & Pacific	High income
...	...	...	...	...	...
443	MacBook Air 13.3" Laptop	$799	United States	North America	High income
444	Apple Pencil (2nd generation)	$129	United States	North America	High income
445	24-inch iMac²	$1,299	United States	North America	High income
446	MacBook Air	$99	United States	North America	High income
447	Magic Mouse	$79	United States	North America	High income

448 rows × 5 columns

3 NF

As the column is is_Amazon_Seller not fully depend on the prod_name it is partially dependent hence we can remove it .

```
In [119]: amzon_2NF= amazon_df.drop(['Uniq_Id','Brand_Name','Model_Number','Shipping_Weight','Selling_Price'], axis=1)

In [124]: amazon_p = amzon_2NF.drop(['Is_Amazon_Seller'],axis = 1)

In [125]: amazon_p.to_sql(name = 'amazon_product', con = engine, if_exists='append', index = False)

Out[125]: 10002

In [ ]:
```

Product_Name	Category	Product_Specification	Product_Url
Rubie's Child's Pokemon Deluxe Pikachu Costum...	Clothing, Shoes & Jewelry	Costumes & Access...	ProductDimensions:16x8x1inches	ItemWeight:1...	https://www.amazon.com/Rubies-Deluxe-Pikac...
Hoffmaster 120813 Double-Tipped Triangular Cr...	Toys & Games	Arts & Crafts	Drawing & Painti...	NULL	https://www.amazon.com/Hoffmaster-120813-...
ARTSCAPE Etched Glass 24" x 36" Window Film,...	Home & Kitchen	Home Décor	Window Treatm...	ProductDimensions:72x36x0inches|ItemWeight:...	https://www.amazon.com/ARTSCAPE-Etched-G...
Pokemon TCG: Sun and Moon Crimson Invasion ...	NULL	ProductDimensions:7.5x3.5x6.8inches|ItemWei...	https://www.amazon.com/Pokemon-TCG-Crims...
Moonlite, 3-Story Bundle for Boys, Includes 3 St...	Toys & Games	Baby & Toddler Toys	ProductDimensions:0.8x5x8.2inches	ItemWeigh...	https://www.amazon.com/Moonlite-3-Story-Bu...
DC Cover Girls: Black Canary by Joëlle Jones St...	Toys & Games	Collectible Toys	Statues, Bobb...	ProductDimensions:4x3x9.4inches|ItemWeight:...	https://www.amazon.com/DC-Cover-Girls-Cana...

amazon_product 30

29988 rows × 6 columns

```
In [139]: walmart_prod = walmart_p.drop(['Available'],axis =1)

In [140]: walmart_prod.to_sql(name = 'walmart_product', con = engine, if_exists='append', index = False)

Out[140]: 29988

In [141]: walmart_prod

Out[141]:
```

	Product_Url	Product_Name	Brand	Product_inovation	Category
0	https://www.walmart.com/ip/La-Costena-Chipotle...	La Costena Chipotle Peppers, 7 OZ (Pack of 12)	La CosteÃ¯Â¿Â½Ã¯Â¿Â½a	NaN	Food	Meal Solutions, Grains & Pasta	Canned...
1	https://www.walmart.com/ip/Equate-Triamcinolon...	Equate Triamcinolone Acetonide Nasal Allergy S...	Equate	NaN	Health	Equate	Equate Allergy	Equate Sinu...
2	https://www.walmart.com/ip/AduroSmart-ERIA-Sof...	AduroSmart ERIA Soft White Smart A19 Light Bul...	AduroSmart ERIA	NaN	Electronics	Smart Home	Smart Energy and Li...
3	https://www.walmart.com/ip/24-Classic-Adjustab...	24" Classic Adjustable Balloon Fender Set Chro...	lowrider	NaN	Sports & Outdoors	Bikes	Bike Accessories	...
4	https://www.walmart.com/ip/Elephant-Shape-Sili...	Elephant Shape Silicone Drinkware Portable Sil...	Anself	NaN	Baby	Feeding	Sippy Cups: Alternatives to P...
...	...	...	...	...	...
29983	https://www.walmart.com/ip/McCain-Smiles/51726947	McCain Smiles	McCain	NaN	Food	Frozen Foods	Frozen Vegetables
29984	https://www.walmart.com/ip/Shock-Sox-Fork-Seal...	Shock Sox Fork Seal Guards 29-36mm Fork Tube 4...	Shock Sox	NaN	Sports & Outdoors	Bikes	Bike Components	...
29985	https://www.walmart.com/ip/Princes-Gooseberrie...	Princes Gooseberries 300g	Princes	NaN	Food	Meal Solutions, Grains & Pasta	Canned...
29986	https://www.walmart.com/ip/Create-Ion-Grace-3-...	Create Ion Grace 3/4 Inches Straight Hair Iron...	Create Ion	NaN	Beauty	Hair Care	Hair Styling Tools	Flat...
29987	https://www.walmart.com/ip/Green-Bell-Takumino...	Green Bell Takuminowaza Two Way Ear Pick Brass...	Takuminowaza	NaN	Beauty	Here for Every Beauty	Featured Shop...

29988 rows × 5 columns

After 2 NR of apple_prod table we can see that we can further decompose the table into 2 tables

- appleproducts
- appleproductincomegroup

here in the first table all the columns are fully dependent on model and in second table all fields are dependent on country

124
125 ●   select * from appleproductincomegroup

Result Grid	Filter Rows:		Export: Wrap Cell Cont

country	region	incomegroup
Australia	East Asia & Pacific	High income
Australia	East Asia & Pacific	High income
Australia	East Asia & Pacific	High income
Australia	East Asia & Pacific	High income
Australia	East Asia & Pacific	High income
Australia	East Asia & Pacific	High income

walmartdata 39   appleproductincomegroup 40 ×

Output

125 ●   select * from appleproducts

Result Grid	Filter Rows:		Export: Wrap Cell

model	price	country
AirPods Max	A$899	Australia
Apple TV 4K	A$249	Australia
Apple Watch SE	From A$429	Australia
Apple Watch Series 3	From A$299	Australia
Sport Band	A$69	Australia
iPad	From A$499	Australia

walmartdata 41   appleproducts 42 ×

Output

Action Output ▾

#	Time	Action