Comparative Shopping of Hardgoods Tennis Equipment

Project Description/Outline:

This case study allowed our team to analyze data from ecommerce website EBAY, website vendor Wilson and one specialty store Academy Sports. Our group used a dataset with 2019 pricing for Walmart utilizing this dataset as our base items to lookup for our tennis hardgoods equipment comparison. While this dataset has 30,000 records of sample data from Walmart, by focusing on products specifically for tennis our group was able to successfully identify and analyze eight separate tennis hardgoods products.

The datasets used by our group are:

- Kaggle Walmart Product Data CSV: https://www.kaggle.com/promptcloud/walmart-product-data-2019

- Web Scraping for tennis hardgoods products for: EBAY, Academy Sports and Wilson

Research Questions:

- Is it better to buy from the vendor directly or from a retailer?
- Is buying from a specialty store a better bargain, or an ecommerce website retailer?
- What type of availability is there from the vendor and a ecommerce retailer in comparison with that data set from Wal-Mart?

Data Cleansing Process:

We filtered this dataset down to the brands which indicated high profile from the Sports department since sporting SKU's are consistent year to year. By using a dataset from 2019, products would be in stock with these various retailers. Our group specifically avoided any cleaning products, food, holiday or any seasonal items in order to provide an accurate representation of information. We then performed a pivot table to determine which brands had the most SKUs for tennis hardgoods which then allowed us to focus on Under Armour, Rawlings, Wilson and Spaulding. Wilson was then determined to have the widest selection of hardgoods products showing availability among these specific retailers.

Since this Kaggle dataset already included Walmart URLs, our team checked to ensure these tennis hardgoods URLs were all still active. We then researched the other competitor's website to find these same URLs for these same products. Our group decided to compare prices of

tennis hardgoods against prices at EBAY (e-commerce website) against both Academy Sports (specialty sporting goods store), as well as directly from Wilson (vendor).

This cleansing process left us with 8 random tennis products from Wilson. Our group also read in Wilson's CSV file.

Product Name

Product Name
Wilson Roger Federer Junior Tennis Racquet
Wilson Blade Team Tennis Racquet
Wilson Burn 100ULS Tennis Racquet
Wilson Championship Extra-Duty Tennis Balls
Wilson Team One Tennis Bag
Wilson Overgrips 3-Pack White
Wilson Overgrips 3-Pack Black
Wilson Women's Rush Pro 2.5 Tennis Shoes

	product_name	category_name	store_name	store_url	url
0	Wilson Roger Federer Junior Tennis Racquet	Racquets	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/rackets/ro...
1	Wilson Blade Team Tennis Racquet	Racquets	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/rackets/bl...
2	Wilson Burn 100ULS Tennis Racquet	Racquets	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/rackets/bu...
3	Wilson Championship Extra-Duty Tennis Balls	Balls	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/balls/valu...
4	Wilson Team One Tennis Bag	Bags	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/bags/tour/...

From this sheet, we added the product category, the individual store URLs, and the store name.

product_name	category_name	store_name	store_url	url
Wilson Roger Federer Junior Tennis Racquet	Racquets	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/rackets/roger-federer-23-tennis-racket-158955
Wilson Blade Team Tennis Racquet	Racquets	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/rackets/blade-team-tennis-racket-156242
Wilson Burn 100ULS Tennis Racquet	Racquets	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/rackets/burn-100s-v4-tennis-racket
Wilson Championship Extra-Duty Tennis Balls	Balls	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/balls/value/championship-extra-duty-tennis-balls-4-ball-can
Wilson Team One Tennis Bag	Bags	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/bags/tour/small-super-tour-2-tennis-bag
Wilson Overgrips 3-Pack White	Accessories	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/grips/pro-overgrip-3-pack-159468
Wilson Overgrips 3-Pack Black	Accessories	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/grips/overgrips/pro-overgrip-sensation-3-pack
Wilson Women's Rush Pro 2.5 Tennis Shoes	Shoes	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/shoes/womens/women-s-rush-pro-3-0-159129
Wilson Roger Federer Junior Tennis Racquet	Racquets	Ebay	https://www.ebay.com/i/154201803162?chn=ps&mkevt=1&mkcid=28	
Wilson Blade Team Tennis Racquet	Racquets	Ebay	https://www.ebay.com/itm/Wilson-Blade-Team-Tennis-Racquet-with-Cover-Brand-New/293311800614?hash=item444abee526:g:a9wAAOSwU51dvzoG	
Wilson Burn 100ULS Tennis Racquet	Racquets	Ebay	https://www.ebay.com/itm/New-Wilson-Burn-100-ULS-Ultra-Lite-Tennis-Racquet-4-1-4-Free-Shipping/274523313950	
Wilson Championship Extra-Duty Tennis Balls	Balls	Ebay	https://www.ebay.com/itm/Wilson-Championship-Extra-Duty-Tennis-Balls-Clive-Brunskill-Sealed-3-Balls-New/143844577471	
Wilson Team One Tennis Bag	Bags	Ebay	https://www.ebay.com/itm/Wilson-Classic-Tennis-Racket-Bag-With-Accessory-Storage-Free-Shipping/284114181272	
Wilson Overgrips 3-Pack White	Accessories	Ebay	https://www.ebay.com/itm/NEW-Wilson-Pro-Overgrip-3-Pack-Tennis-Grip-Tape-Roger-Federer-RF-WHITE/123252083949	
Wilson Overgrips 3-Pack Black	Accessories	Ebay	https://www.ebay.com/itm/Wilson-Pro-Sensation-3-pack-tennis-overgrip/293605926579	
Wilson Women's Rush Pro 2.5 Tennis Shoes	Shoes	Ebay	https://www.ebay.com/itm/Wilson-Rush-Pro-2-5-Women-s-Tennis-Shoes-ASTRAL-AURA/124218162269	
Wilson Roger Federer Junior Tennis Racquet	Racquets	Academy	https://www.academy.com	https://www.academy.com/shop/pdp/wilson-roger-federer-junior-tennis-racquet#repChildCatSku=1
Wilson Blade Team Tennis Racquet	Racquets	Academy	https://www.academy.com	https://www.academy.com/shop/pdp/wilson-blade-team-tennis-racquet#repChildCatid=6779776
Wilson Burn 100ULS Tennis Racquet	Racquets	Academy	https://www.academy.com	https://www.academy.com/shop/pdp/wilson-burn-100uls-tennis-racquet-118179392#repChildCatid
Wilson Championship Extra-Duty Tennis Balls	Balls	Academy	https://www.academy.com	https://www.academy.com/shop/pdp/wilson-championship-extra-duty-tennis-balls-1-can-3-pack#r
Wilson Team One Tennis Bag	Bags	Academy	https://www.academy.com	https://www.academy.com/shop/pdp/wilson-team-1-compartment-small-tennis-bag#repChildCatid
Wilson Overgrips 3-Pack White	Accessories	Academy	https://www.academy.com	https://www.academy.com/shop/pdp/wilson-overgrips-3-pack#repChildCatSku=007724172
Wilson Overgrips 3-Pack Black	Accessories	Academy	https://www.academy.com	https://www.academy.com/shop/pdp/wilson-profile-overgrip-3-pack#repChildCatid=1448186
Wilson Women's Rush Pro 2.5 Tennis Shoes	Shoes	Academy	https://www.academy.com	https://www.academy.com/shop/pdp/wilson-womens-rush-pro-25-tennis-shoes-116129674#repCh

Our group initially had planned on using Dicks Sporting Goods but this specialty retailer has a timing issue and continued to time out when trying to access this store's data online. We

researched pulling products from Amazon but then our team wasn't able to successfully use web scraping from Amazon with our Jupyter's notebook coding. Our group then decided to focus on analyzing data from web scraping from Wilson, Academy Sports and EBAY comparing this information to our Kaggle.com 2019 Wal-Mart dataset for these same eight products which were pulled from three different tennis hardgoods stores.

Once this data was created, we were able to create 2 of our 3 tables in Postgres:
lk_stores
lk_category

LK Category (Item.CSV):
Our group created a data frame using arrays for category name under the headings of

- **Racquets**
- **Balls**
- **Bags**
- **Accessories**
- **Shoes**

LK Store (LK_store.CSV):
The data frame for the stores is grouped by store name and the respective store URL.

	id	store_name	store_url	last_updated
0	13	Academy	https://www.academy.com	11:31.7
1	14	Ebay	https://www.ebay.com	11:31.7
2	15	Wilson	https://www.wilson.com	11:31.7

In order to calculate the product price, our group read in the LK Store CSV file which helped us to analyze the availability (true vs. false) for the selected tennis products. This information was then cross referenced for the selected category names propagating the respective information for each of Wilson's tennis hardgoods products.

	product_name	list_price	scraped_date	on_sale	url	availability
0	Roger Federer 23 Kids Tennis Racket	30.00	2020-12-21 18:26:39.918621	False	https://www.wilson.com/en-us/tennis/rackets/ro...	True
1	Blade Team Tennis Racket	120.00	2020-12-21 18:26:42.514782	False	https://www.wilson.com/en-us/tennis/rackets/bl...	False
2	Burn 100S v4 Tennis Racket	149.00	2020-12-21 18:26:42.990142	False	https://www.wilson.com/en-us/tennis/rackets/bu...	True
3	Championship Extra Duty Tennis Balls - 4 Ball Can	5.99	2020-12-21 18:26:43.347287	False	https://www.wilson.com/en-us/tennis/balls/valu...	False
4	Super Tour 2 Compartment Small Tennis Bag	89.00	2020-12-21 18:26:43.727010	False	https://www.wilson.com/en-us/tennis/bags/tour/...	False

	product_name	category_name	store_name	store_url	url
0	Wilson Roger Federer Junior Tennis Racquet	Racquets	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/rackets/ro...
1	Wilson Blade Team Tennis Racquet	Racquets	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/rackets/bl...
2	Wilson Burn 100ULS Tennis Racquet	Racquets	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/rackets/bu...
3	Wilson Championship Extra-Duty Tennis Balls	Balls	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/balls/valu...
4	Wilson Team One Tennis Bag	Bags	Wilson	https://www.wilson.com	https://www.wilson.com/en-us/tennis/bags/tour/...

Our group scrapped data from each of these three website's for:

- each store's URL
- each store's name
- merging stores with each tennis hardgoods item

```python
url_merge = pd.merge(scraped_data,
                     urls,
                     on ='url',
                     how ='left')
url_merge
```

	product_name_x	list_price	scraped_date	on_sale	url	availability	product_name_y	category_name	store_name
0	Roger Federer 23 Kids Tennis Racket	30.00	2020-12-21 18:26:39.918621	False	https://www.wilson.com/en-us/tennis/rackets/ro...	True	Wilson Roger Federer Junior Tennis Racquet	Racquets	Wilson
1	Blade Team Tennis Racket	120.00	2020-12-21 18:26:42.514782	False	https://www.wilson.com/en-us/tennis/rackets/bl...	False	Wilson Blade Team Tennis Racquet	Racquets	Wilson
2	Burn 100S v4 Tennis Racket	149.00	2020-12-21 18:26:42.990142	False	https://www.wilson.com/en-us/tennis/rackets/bu...	True	Wilson Burn 100ULS Tennis Racquet	Racquets	Wilson
3	Championship Extra Duty Tennis Balls - 4 Ball Can	5.99	2020-12-21 18:26:43.347287	False	https://www.wilson.com/en-us/tennis/balls/valu...	False	Wilson Championship Extra-Duty Tennis Balls	Balls	Wilson
4	Super Tour 2 Compartment Small Tennis Bag	89.00	2020-12-21 18:26:43.727010	False	https://www.wilson.com/en-us/tennis/bags/tour/...	False	Wilson Team One Tennis Bag	Bags	Wilson
5	Pro Overgrip, 3 Pack	6.49	2020-12-21 18:26:44.069455	False	https://www.wilson.com/en-us/tennis/grips/pro-...	True	Wilson Overgrips 3-Pack White	Accessories	Wilson
6	Pro Overgrip Sensation, 3 Pack	6.49	2020-12-21 18:26:44.513761	False	https://www.wilson.com/en-us/tennis/grips/over...	True	Wilson Overgrips 3-Pack Black	Accessories	Wilson
7	Women's Rush Pro 3.0 Tennis Shoe	129.00	2020-12-21 18:26:45.007701	True	https://www.wilson.com/en-us/tennis/shoes/wome...	True	Wilson Women's Rush Pro 2.5 Tennis Shoes	Shoes	Wilson
8	Wilson Roger Federer Junior Tennis Racquet	24.99	2020-12-21 18:26:53.034991	False	https://www.academy.com/shop/pdp/wilson-roger-...	True	Wilson Roger Federer Junior Tennis Racquet	Racquets	Academy
9	Wilson Blade Team Tennis Racquet	119.99	2020-12-21 18:26:56.954163	False	https://www.academy.com/shop/pdp/wilson-blade-...	True	Wilson Blade Team Tennis Racquet	Racquets	Academy
10	Wilson Burn 100ULS Tennis Racquet	139.99	2020-12-21 18:27:01.103138	False	https://www.academy.com/shop/pdp/wilson-burn-1...	True	Wilson Burn 100ULS Tennis Racquet	Racquets	Academy

11	Wilson Championship Extra-Duty Tennis Balls 1 ...	2.49	2020-12-21 18:27:08.129922	False	https://www.academy.com/shop/pdp/wilson-champi...	True	Wilson Championship Extra-Duty Tennis Balls	Balls	Academy
12	Wilson Team 1 Compartment Small Tennis Bag	39.99	2020-12-21 18:27:14.143755	False	https://www.academy.com/shop/pdp/wilson-team-1...	False	Wilson Team One Tennis Bag	Bags	Academy
13	Wilson Overgrips 3-Pack	5.99	2020-12-21 18:27:19.667225	False	https://www.academy.com/shop/pdp/wilson-overgr...	True	Wilson Overgrips 3-Pack White	Accessories	Academy
14	Wilson Profile Overgrip 3-Pack	5.99	2020-12-21 18:27:25.070174	False	https://www.academy.com/shop/pdp/wilson-profil...	True	Wilson Overgrips 3-Pack Black	Accessories	Academy
15	Wilson Women's Rush Pro 2.5 Tennis Shoes	129.99	2020-12-21 18:27:29.869464	False	https://www.academy.com/shop/pdp/wilson-womens...	False	Wilson Women's Rush Pro 2.5 Tennis Shoes	Shoes	Academy
16	Details about Wilson Youth Roger Federer Rec...	28.13	2020-12-21 18:27:33.932244	False	https://www.ebay.com/itm/154201803162?chn=ps&m...	True	Wilson Roger Federer Junior Tennis Racquet	Racquets	Ebay
17	Details about Wilson Blade Team Tennis Racqu...	38.99	2020-12-21 18:27:37.527766	False	https://www.ebay.com/itm/Wilson-Blade-Team-Ten...	True	Wilson Blade Team Tennis Racquet	Racquets	Ebay
18	Details about New Wilson Burn 100 ULS Ultra ...	107.00	2020-12-21 18:27:40.818746	False	https://www.ebay.com/itm/New-Wilson-Burn-100-U...	True	Wilson Burn 100ULS Tennis Racquet	Racquets	Ebay
19	Details about Wilson Championship Extra Duty...	24.99	2020-12-21 18:27:45.060829	False	https://www.ebay.com/itm/Wilson-Championship-E...	True	Wilson Championship Extra-Duty Tennis Balls	Balls	Ebay
20	Details about Wilson Classic Tennis Racket B...	35.00	2020-12-21 18:27:48.331804	False	https://www.ebay.com/itm/Wilson-Classic-Tennis...	True	Wilson Team One Tennis Bag	Bags	Ebay
21	Details about NEW Wilson Pro Overgrip - 3 Pa...	9.99	2020-12-21 18:27:51.982639	False	https://www.ebay.com/itm/NEW-Wilson-Pro-Overgr...	True	Wilson Overgrips 3-Pack White	Accessories	Ebay
22	Details about Wilson Pro Sensation 3-pack te...	6.49	2020-12-21 18:27:54.854563	False	https://www.ebay.com/itm/Wilson-Pro-Sensation-...	True	Wilson Overgrips 3-Pack Black	Accessories	Ebay
23	Details about Wilson Rush Pro 2.5 Women's Te...	65.00	2020-12-21 18:27:57.496628	False	https://www.ebay.com/itm/Wilson-Rush-Pro-2-5-W...	True	Wilson Women's Rush Pro 2.5 Tennis Shoes	Shoes	Ebay

```python
store_merge = pd.merge(url_merge,
                       stores,
                       on ='store_name',
                       how ='left')
store_merge.head()
```

	product_name_x	list_price	scraped_date	on_sale	url	availability	product_name_y	category_name	store_name
0	Roger Federer 23 Kids Tennis Racket	30.00	2020-12-21 18:26:39.918621	False	https://www.wilson.com/en-us/tennis/rackets/ro...	True	Wilson Roger Federer Junior Tennis Racquet	Racquets	Wilson
1	Blade Team Tennis Racket	120.00	2020-12-21 18:26:42.514782	False	https://www.wilson.com/en-us/tennis/rackets/bl...	False	Wilson Blade Team Tennis Racquet	Racquets	Wilson
2	Burn 100S v4 Tennis Racket	149.00	2020-12-21 18:26:42.990142	False	https://www.wilson.com/en-us/tennis/rackets/bu...	True	Wilson Burn 100ULS Tennis Racquet	Racquets	Wilson
3	Championship Extra Duty Tennis Balls - 4 Ball Can	5.99	2020-12-21 18:26:43.347287	False	https://www.wilson.com/en-us/tennis/balls/valu...	False	Wilson Championship Extra-Duty Tennis Balls	Balls	Wilson
4	Super Tour 2 Compartment Small Tennis Bag	89.00	2020-12-21 18:26:43.727010	False	https://www.wilson.com/en-us/tennis/bags/tour/...	False	Wilson Team One Tennis Bag	Bags	Wilson

```
category_merge = pd.merge(store_merge,
                          items,
                          on ='category_name',
                          how ='left')
category_merge.head()
```

	product_name_x	list_price	scraped_date	on_sale	url	availability	product_name_y	category_name	store_name
0	Roger Federer 23 Kids Tennis Racket	30.00	2020-12-21 18:26:39.918621	False	https://www.wilson.com/en-us/tennis/rackets/ro...	True	Wilson Roger Federer Junior Tennis Racquet	Racquets	Wilson
1	Blade Team Tennis Racket	120.00	2020-12-21 18:26:42.514782	False	https://www.wilson.com/en-us/tennis/rackets/bl...	False	Wilson Blade Team Tennis Racquet	Racquets	Wilson
2	Burn 100S v4 Tennis Racket	149.00	2020-12-21 18:26:42.990142	False	https://www.wilson.com/en-us/tennis/rackets/bu...	True	Wilson Burn 100ULS Tennis Racquet	Racquets	Wilson
3	Championship Extra Duty Tennis Balls - 4 Ball Can	5.99	2020-12-21 18:26:43.347287	False	https://www.wilson.com/en-us/tennis/balls/valu...	False	Wilson Championship Extra-Duty Tennis Balls	Balls	Wilson
4	Super Tour 2 Compartment Small Tennis Bag	89.00	2020-12-21 18:26:43.727010	False	https://www.wilson.com/en-us/tennis/bags/tour/...	False	Wilson Team One Tennis Bag	Bags	Wilson

store_url_x	id_x	store_url_y	last_updated_x	id_y	last_updated_y
https://www.wilson.com	15	https://www.wilson.com	11:31.7	6	20:22.4
https://www.wilson.com	15	https://www.wilson.com	11:31.7	6	20:22.4
https://www.wilson.com	15	https://www.wilson.com	11:31.7	6	20:22.4
https://www.wilson.com	15	https://www.wilson.com	11:31.7	7	20:22.4
https://www.wilson.com	15	https://www.wilson.com	11:31.7	8	20:22.4

In order to merge the categories, we renamed the columns accordingly and then printed
Product_Price to show the scrapped website data. We did notice that there were not as many
'on sale' price points so we ended up deleting this information.

```
category_merge.rename(columns={'product_name_x':'product_name'}, inplace=True)
category_merge.rename(columns={'store_url_x':'store_url'}, inplace=True)
category_merge.rename(columns={'id_x' : 'store_id'}, inplace=True)
category_merge.rename(columns={'id_y' : 'category_id'}, inplace=True)
```

```
product_price = category_merge[["product_name","category_id","store_id","list_price","scraped_date","on_sale","url"]]
product_price.set_index('product_name', inplace=True)
product_price
```

product_name	category_id	store_id	list_price	scraped_date	on_sale	url
Roger Federer 23 Kids Tennis Racket	6	15	30.00	2020-12-21 18:26:39.918621	False	https://www.wilson.com/en-us/tennis/rackets/ro...
Blade Team Tennis Racket	6	15	120.00	2020-12-21 18:26:42.514782	False	https://www.wilson.com/en-us/tennis/rackets/bl...
Burn 100S v4 Tennis Racket	6	15	149.00	2020-12-21 18:26:42.990142	False	https://www.wilson.com/en-us/tennis/rackets/bu...
Championship Extra Duty Tennis Balls - 4 Ball Can	7	15	5.99	2020-12-21 18:26:43.347287	False	https://www.wilson.com/en-us/tennis/balls/valu...
Super Tour 2 Compartment Small Tennis Bag	8	15	89.00	2020-12-21 18:26:43.727010	False	https://www.wilson.com/en-us/tennis/bags/tour/...
Pro Overgrip, 3 Pack	9	15	6.49	2020-12-21 18:26:44.069455	False	https://www.wilson.com/en-us/tennis/grips/pro-...
Pro Overgrip Sensation, 3 Pack	9	15	6.49	2020-12-21 18:26:44.513761	False	https://www.wilson.com/en-us/tennis/grips/over...
Women's Rush Pro 3.0 Tennis Shoe	10	15	129.00	2020-12-21 18:26:45.007701	True	https://www.wilson.com/en-us/tennis/shoes/wome...
Wilson Roger Federer Junior Tennis Racquet	6	13	24.99	2020-12-21 18:26:53.034991	False	https://www.academy.com/shop/pdp/wilson-roger-...
Wilson Blade Team Tennis Racquet	6	13	119.99	2020-12-21 18:26:56.954163	False	https://www.academy.com/shop/pdp/wilson-blade-...
Wilson Burn 100ULS Tennis Racquet	6	13	139.99	2020-12-21 18:27:01.103138	False	https://www.academy.com/shop/pdp/wilson-burn-1...
Wilson Championship Extra-Duty Tennis Balls 1 Can/3-Pack	7	13	2.49	2020-12-21 18:27:08.129922	False	https://www.academy.com/shop/pdp/wilson-champi...
Wilson Team 1 Compartment Small Tennis Bag	8	13	39.99	2020-12-21 18:27:14.143755	False	https://www.academy.com/shop/pdp/wilson-team-1...
Wilson Overgrips 3-Pack	9	13	5.99	2020-12-21 18:27:19.667225	False	https://www.academy.com/shop/pdp/wilson-overgr...
Wilson Profile Overgrip 3-Pack	9	13	5.99	2020-12-21 18:27:25.070174	False	https://www.academy.com/shop/pdp/wilson-profil...
Wilson Profile Overgrip 3-Pack	9	13	5.99	2020-12-21 18:27:25.070174	False	https://www.academy.com/shop/pdp/wilson-profil...
Wilson Women's Rush Pro 2.5 Tennis Shoes	10	13	129.99	2020-12-21 18:27:29.869464	False	https://www.academy.com/shop/pdp/wilson-womens...
Details about Wilson Youth Roger Federer Recreational Tennis Racket - 23"	6	14	28.13	2020-12-21 18:27:33.932244	False	https://www.ebay.com/itm/154201803162?chn=ps&m...
Details about Wilson Blade Team Tennis Racquet with Cover - (Brand New!)	6	14	38.99	2020-12-21 18:27:37.527766	False	https://www.ebay.com/itm/Wilson-Blade-Team-Ten...
Details about New Wilson Burn 100 ULS Ultra Lite Tennis Racquet, 4 1/4", Free Shipping	6	14	107.00	2020-12-21 18:27:40.818746	False	https://www.ebay.com/itm/New-Wilson-Burn-100-U...
Details about Wilson Championship Extra Duty Tennis Balls Clive Brunskill Sealed 3 Balls New	7	14	24.99	2020-12-21 18:27:45.060829	False	https://www.ebay.com/itm/Wilson-Championship-E...
Details about Wilson Classic Tennis Racket Bag With Accessory Storage Free Shipping	8	14	35.00	2020-12-21 18:27:48.331804	False	https://www.ebay.com/itm/Wilson-Classic-Tennis...
Details about NEW Wilson Pro Overgrip - 3 Pack Tennis Grip Tape Roger Federer RF WHITE	9	14	9.99	2020-12-21 18:27:51.982639	False	https://www.ebay.com/itm/NEW-Wilson-Pro-Overgr...
Details about Wilson Pro Sensation 3-pack tennis overgrip	9	14	6.49	2020-12-21 18:27:54.854563	False	https://www.ebay.com/itm/Wilson-Pro-Sensation-...
Details about Wilson Rush Pro 2.5 Women's Tennis Shoes ASTRAL AURA	10	14	65.00	2020-12-21 18:27:57.496628	False	https://www.ebay.com/itm/Wilson-Rush-Pro-2-5-W...

The URL's we scrapped from Wilson's, Academy Sports and EBAY's websites are:

```
#Wilson item urls
wRac1 = "rackets/roger-federer-23-tennis-racket-158955"
wRac2 = "rackets/blade-team-tennis-racket-156242"
wRac3 = "rackets/burn-100s-v4-tennis-racket"
wBall = "balls/value/championship-extra-duty-tennis-balls-4-ball-can"
wBag = "bags/tour/small-super-tour-2-tennis-bag"
wAcc1 = "grips/pro-overgrip-3-pack-159468"
wAcc2 = "grips/overgrips/pro-overgrip-sensation-3-pack"
wShoe = "shoes/womens/women-s-rush-pro-3-0-159129"
```

```
#Academy item urls
aRac1 = "wilson-roger-federer-junior-tennis-racquet#repChildCatSku=122845378"
aRac2 = "wilson-blade-team-tennis-racquet#repChildCatid=6779776"
aRac3 = "wilson-burn-100uls-tennis-racquet-118179392#repChildCatid=6759657"
aBall = "wilson-championship-extra-duty-tennis-balls-1-can-3-pack#repChildCatid=460610"
aBag = "wilson-team-1-compartment-small-tennis-bag#repChildCatid=6779780"
aAcc1 = "wilson-overgrips-3-pack#repChildCatSku=007724172"
aAcc2 = "wilson-profile-overgrip-3-pack#repChildCatid=1448186"
aShoe = "wilson-womens-rush-pro-25-tennis-shoes-116129674#repChildCatid=5262044"
```

```
#Ebay item urls
eRac1 = "itm/154201803162?chn=ps&mkevt=1&mkcid=28"
eRac2 = "itm/Wilson-Blade-Team-Tennis-Racquet-with-Cover-Brand-New/293311800614?hash=item444abee526:g:a9wAAOSwU51dvzoG"
eRac3 = "itm/New-Wilson-Burn-100-ULS-Ultra-Lite-Tennis-Racquet-4-1-4-Free-Shipping/274523313950"
eBall = "itm/Wilson-Championship-Extra-Duty-Tennis-Balls-Clive-Brunskill-Sealed-3-Balls-New/143844577471"
eBag = "itm/Wilson-Classic-Tennis-Racket-Bag-With-Accessory-Storage-Free-Shipping/284114181272"
eAcc1 = "itm/NEW-Wilson-Pro-Overgrip-3-Pack-Tennis-Grip-Tape-Roger-Federer-RF-WHITE/123252083949"
eAcc2 = "itm/Wilson-Pro-Sensation-3-pack-tennis-overgrip/293605926579"
eShoe = "itm/Wilson-Rush-Pro-2-5-Women-s-Tennis-Shoes-ASTRAL-AURA/124218162269"
```

**For web scraping, we used both Beautiful Soup and Splinter.

An example of the code written for the web scrapping process for Wilson's is shown for example. We followed this same logic to complete the web scrapping for Academy Sports and EBAY's websites.

```python
#masterDF = pd.DataFrame()

wProductsWeWant = [wRac1, wRac2, wRac3, wBall, wBag, wAcc1, wAcc2, wShoe]

productNames = []
listPrices = []
onSales = []
urls = []
availabilitys = []
scrapedDates = []


for wProductToScrape in wProductsWeWant:

    #get correct URL format for the products
    urlW = f"https://www.wilson.com/en-us/tennis/{wProductToScrape}"
    response = requests.get(urlW)

    soup = bs(response.content, "lxml")

    #pull data, need to find where data is located
    productName = soup.find("div", class_="product-name").text.strip()
    listPrice = soup.find("span", class_="price").text.strip()
    url1 = urlW

    if soup.find("span", class_="old-price"):
        onSale = True
    else:
        onSale = False

    if soup.find("span", class_="button-title").text == 'Out of Stock':
        availability=False
    else:
        availability=True

    scrapedDate = pd.to_datetime('today')
```

```python
    #appending data
    productNames.append(productName)
    listPrices.append(listPrice)
    onSales.append(onSale)
    urls.append(url1)
    availabilitys.append(availability)
    scrapedDates.append(scrapedDate)

#create DF
wDF=pd.DataFrame()
wDF["product_name"] = productNames
wDF["list_price"] = listPrices
wDF["scraped_date"] = scrapedDates
wDF["on_sale"] = onSales
wDF["url"] = urls
wDF["availability"] = availabilitys
wDF.set_index('product_name', inplace=True)
wDF["list_price"] = wDF["list_price"].replace('[$]', '', regex=True).astype('float')

#add info to DF
masterDF = pd.DataFrame(wDF)
```

This data was then imported into SQL tables using Jupyter Notebook. After all of our code was completed, then this data was formatted into pgAdmin4. PgAdmin4 is used to house our tables. As a side note, our group is using a relational database since we are using SQL, in turn helping to results in a greater variety of indexing capabilities resulting in faster query response times.

We created a table using SQL for Product_Price, Lk_Category and LK_Store using a serial primary key to create unique values to create columns for our tables. Additionally, Varchar (100) allowed us to create a string to contain information for the category name.

```sql
CREATE TABLE "product_price" (
    id serial primary key,
    product_name varchar(100) NOT NULL,
    category_id int NOT NULL,
    store_id int NOT NULL,
    list_price float NOT NULL,
    scraped_date timestamp NOT NULL,
    on_sale boolean NOT NULL,
    url varchar(500) NOT NULL,
    availability boolean NOT NULL,
        last_updated timestamp default localtimestamp,
        foreign key (category_id) REFERENCES "lk_category" ("id"),
        foreign key (store_id)REFERENCES "lk_store" ("id")
);

CREATE TABLE "lk_category" (
    id serial primary key,
    category_name varchar(100) NOT NULL,
    last_updated timestamp default localtimestamp
);

CREATE TABLE "lk_store" (
    id serial primary key,
    store_name varchar(100) NOT NULL,
    store_url varchar(100) NOT NULL,
    last_updated timestamp default localtimestamp
);
```

Project Limitations:

- Sale price may differ or not be a relative factor between one retailer vs. another retailer (i.e., Wal-Mart chooses to have their tennis hardgoods products on sale but Wilson chooses to have their tennis hardgoods product $1-2 more since Wilson is a vendor retailer for this industry)

- Wal-Mart's website is secured by an API so our group isn't able to be utilize web scraping to obtain information from Wal-Mart.com. However, our group was able to use web scraping for Wilson and Academy Sports.

- Needed to delete sale price since there wasn't enough data for sale price to use for retailer analysis

ERD Diagram:

```
 1   product_price
 2   -
 3   id PK int
 4   product-name varchar(100)
 5   category_id int FK >- lk_category.id
 6   store_id int FK >- lk_store.id
 7   list_price int
 8   scraped_date int
 9   on_sale boolean
10   url varchar(500)
11   last_updated datetime
12
13   lk_category
14   -
15   id PK int
16   cateogry_name varchar(100)
17   last_updated datetime
18
19   lk_store
20   -
21   id PK int
22   store_name varchar(100)
23   store_url varchar(100)
24   availability boolean
25   last_updated datetime
```

lk_category

id	int
cateogry_name	varchar(100)
last_updated	datetime

product_price

id	int
product-name	varchar(100)
category_id	int
store_id	int
list_price	int
scraped_date	int
on_sale	boolean
url	varchar(500)
last_updated	datetime

lk_store

id	int
store_name	varchar(100)
store_url	varchar(100)
availability	boolean
last_updated	datetime

Query Screen Shots from PgAdmin4:

postgres/postgres@etl_project ∨

Query Editor   Query History

```sql
1  Select
2  ct.category_name,
3  avg(pr.list_price) as avg_price
4  From
5  product_price pr
6  join lk_category ct on pr.category_id=ct.id
7  group by ct.category_name
8  order by avg_price desc
9  ;
```

Data Output   Explain   Messages   Notifications

	category_name character varying (100)	avg_price double precision
1	Shoes	107.99666666666667
2	Racquets	84.23222222222222
3	Bags	54.663333333333334
4	Balls	11.156666666666666
5	Accessories	6.906666666666667

postgres/postgres@etl_project ∨

Query Editor   Query History

```sql
1  Select
2  st.store_name,
3  avg(pr.list_price) as avg_price
4  From
5  product_price pr
6  join lk_store st on pr.store_id=st.id
7  group by st.store_name
8  order by avg_price desc
```

Data Output   Explain   Messages   Notifications

	store_name character varying (100)	avg_price double precision
1	Wilson	66.99625
2	Academy	58.67750000000001
3	Ebay	39.448750000000004

```
1   Select
2   st.store_name,
3   ct.category_name,
4   avg(pr.list_price) as avg_price
5   From
6   product_price pr
7   join lk_store st on pr.store_id=st.id
8   join lk_category ct on pr.category_id=ct.id
9   group by st.store_name, ct.category_name
10  order by avg_price desc
```

Data Output   Explain   Messages   Notifications

	store_name character varying (100) 🔒	category_name character varying (100) 🔒	avg_price double precision 🔒
1	Academy	Shoes	129.99
2	Wilson	Shoes	129
3	Wilson	Racquets	99.66666666666667
4	Academy	Racquets	94.99000000000001
5	Wilson	Bags	89
6	Ebay	Shoes	65
7	Ebay	Racquets	58.04
8	Academy	Bags	39.99
9	Ebay	Bags	35
10	Ebay	Balls	24.99
11	Ebay	Accessories	8.24
12	Wilson	Accessories	6.49
13	Wilson	Balls	5.99
14	Academy	Accessories	5.99
15	Academy	Balls	2.49

```sql
Select
ct.category_name,
avg(pr.list_price) as avg_price
From
product_price pr
join lk_category ct on pr.category_id=ct.id
group by ct.category_name
order by avg_price desc
;

Select
st.store_name,
avg(pr.list_price) as avg_price
From
product_price pr
join lk_store st on pr.store_id=st.id
group by st.store_name
order by avg_price desc
;

Select
st.store_name,
ct.category_name,
avg(pr.list_price) as avg_price
From
product_price pr
join lk_store st on pr.store_id=st.id
join lk_category ct on pr.category_id=ct.id
group by st.store_name, ct.category_name
order by avg_price desc
;

SELECT
pr.product_name,
pr.list_price,
st.id
FROM
product_price pr
JOIN lk_store st ON pr.store_id=st.id

SELECT
pr.product_name,
pr.availability,
ct.id
FROM
product_price pr
JOIN lk_category ct ON pr.category_id=ct.id
;

SELECT
```

```
50    SELECT
51    pr.product_name,
52    pr.list_price,
53    st.id,
54    st.store_name
55    FROM
56    product_price pr
57    JOIN lk_store st ON pr.store_id=st.id
58    WHERE store_name = 'Wilson'
59    ;
60
61    SELECT
62    SUM(list_price)
63    FROM
64    product_price
65    WHERE
66    availability = 'True'
67    ;
68
69    SELECT
70    pr.product_name,
71    pr.url,
72    ct.category_name
73    FROM
74    product_price pr
75    JOIN lk_category ct ON pr.category_id=ct.id
76    ORDER BY category_name
77    ;
78    }
```

In conclusion, tennis hardgoods products are comparable in price for both Academy Sports compared to Wilson's. However, EBAY's products are usually cheaper in price for these same eight products compared to both Academy Sports and Wilson's. EBAY also has third party resellers and independent sellers who can set their own lowest reseller price in order to be competitive online with other retailers and ecommerce websites.