

Project - Electronic Product Pricing

DESCRIPTION

Problem Statement

- This is a list of over 5000 electronic products with pricing information across 25 unique fields provided by Kaggle. The dataset also includes the brand, category, merchant, name, manufacturer, source, and more.

Goal

- The aim is to understand how each brand, category, manufacturer price has changed over the years
- Is there a correlation between the data points of this dataset
- What is the competitive pricing strategy for the same product from different merchants

Data Snapshot

In [3] :

```
Image ('electronicscreenshot.png')
```

Out [3] :

		id	prices.availability	prices.condition	prices.currency	prices.dateSeen	prices.isSale	prices.merchant	prices.shipping
0		AVphrugr1cnluZ0-FOeH	Yes	New	USD	2017-05-05T00:00:00Z,2017-05-09T15:00:00Z	False	Bestbuy.com	NaN http://
1		AVrl6FDbv8e3D1O-lm4R	Yes	New	USD	2017-10-10T02:00:00Z,2017-08-12T03:00:00Z,2017...	False	Bestbuy.com	NaN https://
2		AVpiLlubilAPnD_xBoTa	Yes	New	USD	2017-10-10T19:00:00Z,2017-09-12T14:00:00Z,2017...	False	Bestbuy.com	NaN https://

Importing All the library

In [4] :

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.preprocessing import LabelEncoder
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression, Ridge, Lasso
from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score
```

```
from statsmodels.formula.api import ols
from sklearn.ensemble import RandomForestRegressor
from sklearn.tree import DecisionTreeRegressor
from sklearn.svm import SVR
from sklearn.decomposition import PCA
import xgboost as xgb
from IPython.display import Image
from sklearn.model_selection import train_test_split
from IPython.display import Image
```

In [5]:

```
pd.set_option('display.max_rows', 100000)
pd.set_option('display.max_columns', 500000)
```

Loading Data Set

In [16]:

```
dt_electronics=pd.read_csv("electronics_products_pricing.csv")
```

In [17]:

```
dt_electronics.head()
```

Out[17]:

		id	prices.availability	prices.condition	prices.currency	
0		AVphrgr1cnluZ0-FOeH	Yes	New	USD	10T
1		AVrl6FDbv8e3D1O-lm4R	Yes	New	USD	10T 12T
2		AVpiLlubilAPnD_xBoTa	Yes	New	USD	10T 12T
3		AVpgQP5vLJeJML43LQbd	Yes	New	USD	08T 18T
4		AV1YDsamoGV-KLJ3adcbe	More on the Way	New	USD	

In [18]:

```
dt_electronics.shape
```

Out[18]:

(5436, 25)

In [19]:

```
dt_electronics.describe()
```

Out[19]:

	price
count	5436.000000
mean	492.941161
std	769.246463
min	1.000000
25%	79.950000
50%	194.410000
75%	486.952500
max	6999.990000

In [20]:

```
dt_electronics.isnull().sum()
```

Out[20]:

id	0
prices.availability	0
prices.condition	0
prices.currency	0
prices.dateSeen	0
prices.isOnSale	0
prices.merchant	0
prices.shipping	2237
prices.sourceURLs	0
asins	0
brand	0
categories	0
dateAdded	0
dateUpdated	0
ean	4261
imageURLs	0
keys	0
manufacturer	2959
manufacturerNumber	0
name	0
primaryCategories	0
sourceURLs	0
upc	0
weight	0
price	0
dtype:	int64

Feature-Engineering

In [21]:

```
dt_electronics.columns
```

Out[21]:

```
Index(['id', 'prices.availability', 'prices.condition', 'prices.currency',
       'prices.dateSeen', 'prices.isSale', 'prices.merchant',
       'prices.shipping', 'prices.sourceURLs', 'asins', 'brand',
       'categories',
       'dateAdded', 'dateUpdated', 'ean', 'imageURLs', 'keys',
       'manufacturer',
       'manufacturerNumber', 'name', 'primaryCategories', 'sourceURLs',
       'upc',
       'weight', 'price'],
      dtype='object')
```

There are some columns with . I am replacing those names

In [22]:

```
dt_electronics.rename(columns={'prices.availability':'pricesavailability',
                             'prices.condition':'pricescondition','prices.currency':'pricescurrency',
                             'prices.dateSeen':'pricesdateSeen','prices.isSale':'pricesisSale','prices.merchant':'pricesmerchant',
                             'prices.shipping':'pricesshipping','prices.sourceURLs':'pricessourceURLs'},inplace=True)
```

In [23]:

```
dt_electronics.head()
```

Out[23]:

		id	pricesavailability	pricescondition	pricescurrency	price
0		AVphrugr1cnluZ0-FOeH	Yes	New	USD	10T20
1		AVrl6FDbv8e3D1O-lm4R	Yes	New	USD	10T0
2		AVpiLlubilAPnD_xBoTa	Yes	New	USD	10T12T14
3		AVpgQP5vLJeJML43LQbd	Yes	New	USD	08T0

4 AV1YDsmoGV-KLJ3adcbe More on the Way

New

USD

I will drop the Id column

In [24]:

```
dt_electronics.drop('id',axis=1,inplace=True)
```

Finding different unique values for each columns

In [25]:

```
for feature in dt_electronics.columns[:]:
    print(feature,":",len(dt_electronics[feature].unique()),'labels')
```

```
pricesavailability : 13 labels
pricescondition : 11 labels
pricescurrency : 2 labels
pricesdateSeen : 3066 labels
pricesisSale : 2 labels
pricesmerchant : 587 labels
pricesshipping : 60 labels
pricessourceURLs : 3562 labels
asins : 819 labels
brand : 254 labels
categories : 792 labels
dateAdded : 784 labels
dateUpdated : 727 labels
ean : 118 labels
imageURLs : 819 labels
keys : 819 labels
manufacturer : 198 labels
manufacturerNumber : 819 labels
name : 813 labels
primaryCategories : 5 labels
sourceURLs : 819 labels
upc : 385 labels
weight : 578 labels
price : 2059 labels
```

Finding the value of each datatypes of this dataset

In [26]:

```
dt_electronics.dtypes
```

Out[26]:

```
pricesavailability      object
pricescondition        object
pricescurrency         object
pricesdateSeen         object
pricesisSale            bool
pricesmerchant          object
pricesshipping          object
pricessourceURLs       object
asins                  object
brand                  object
categories              object
dateAdded               object
dateUpdated              object
ean                     object
imageURLs              object
keys                   object
manufacturer            object
manufacturerNumber      object
name                    object
primaryCategories      object
sourceURLs              object
upc                     object
weight                  object
price                  float64
dtype: object
```

As all the columns are objects we have to convert them so that we can apply any kind of algorithm

working with pricesavailability column

In [27]:

```
dt_electronics.pricesavailability.value_counts()
```

Out[27]:

```
In Stock           2395
Yes                1590
yes                657
TRUE               513
Out Of Stock       89
Special Order      80
More on the Way    63
undefined          28
sold                14
No                 4
FALSE               1
7 available        1
```

```
Retired          1  
Name: pricesavailability, dtype: int64
```

In [28]:

```
dt_electronics.pricesavailability.unique()
```

Out[28]:

```
array(['Yes', 'More on the Way', 'In Stock', 'yes', 'TRUE',  
       'Special Order', 'Out Of Stock', 'undefined', 'No', 's  
old',  
       '7 available', 'FALSE', 'Retired'], dtype=object)
```

I will calculate the count of each unique values of pricesavailability and replace the value with the count

In [29]:

```
pricesavailability_map=dt_electronics['pricesavailability'].value_counts()  
.to_dict()
```

In [30]:

```
pricesavailability_map
```

Out[30]:

```
{'In Stock': 2395,  
'Yes': 1590,  
'yes': 657,  
'TRUE': 513,  
'Out Of Stock': 89,  
'Special Order': 80,  
'More on the Way': 63,  
'undefined': 28,  
'sold': 14,  
'No': 4,  
'FALSE': 1,  
'7 available': 1,  
'Retired': 1}
```

here I will replace 'In Stock' with 2395

In [31]:

```
dt_electronics['pricesavailability']=dt_electronics['pricesavailability'].  
map(pricesavailability_map)  
dt_electronics.head()
```

Out[31]:

pricesavailability	pricescondition	pricescurrency	pricesdateSeen	pricesisSa
--------------------	-----------------	----------------	----------------	------------

0	1590	New	USD	2017-05-05T20:00:00Z,2017-05-09T15:00:00Z	False
1	1590	New	USD	2017-10-10T02:00:00Z,2017-08-12T03:00:00Z,2017...	False
2	1590	New	USD	2017-10-10T19:00:00Z,2017-09-12T14:00:00Z,2017...	False
3	1590	New	USD	2017-09-08T05:00:00Z,2017-09-18T13:00:00Z,2017...	False
4	63	New	USD	2017-12-05T13:00:00Z	True

Now I will apply the same process for pricescondition

In [32]:

```
dt_electronics.pricescondition.unique()
```

Out[32]:

```
array(['New', 'new', 'Refurbished', 'Manufacturer refurbished',
       'Seller refurbished', 'Used', 'New other (see details)',
       'pre-owned',
       '5/16" Ring Terminal, 3 ft. 8 GA Black Ground Cable, 6
       ft. Split Loom Tubing, Depth: 6.5" (165mm) (top) 11.2" (285m
       m) (bottom), Item Weight: 18.5 lbs., Frequency Response 25Hz
       - 500Hz, Line Output, Max Power: 1100 Watt x 1 Channel @ 2 Oh
       m, 30 ft. Speaker Wire, Boss Illuminated Logo, Height: 12.8"
       (325mm), (3) Rubber Grommets, Item Weight: 2 lbs., Size 10",
       20 ft. 8 GA Red Power Cable, Ported enclosure for greater out
       put and deeper bass, 2 Ohm Stable, Class A/B, Voice Coil Size
       2", Black rubber surround, Nominal Impedance 4 ohm, Rugged fi
       berboard construction with thick carpet covering, Warranty: 1
       Year Manufacturer Warranty, MOSFET Power, Weight: 6 lbs, Widt
       h: 17.7" (450mm), Condition: Brand New!, Low Pass Crossover,
       List item, RMS Power: 250 Watt x 1 Channel @ 4 Ohm, Remote Ba
       ss Control Included!, 1/4" Ring Terminal, 16 ft. 18 GA Blue T
       urn-On Wire, Peak Power: 500 Watts, Competition High Quality
       Fuse Holder, Condition: BRAND NEW!, Product Condition: Brand
       New, RMS Power: 175 Watts, Aluminum woofer cone, THD: 0.01%,
       1 Year Manufacturer Warrantv. Dimensions: 10-7/16" (W) x 2-1/
```

```
4" (H) x 9-1/8" (L), #10 Ring Terminal, 20 ft. High Performance Black RCA, SPL (db @ 1w/1m) 88dB, New Dual SBX101 10" 500 Watts Car Audio Subwoofer Sub + Ported Sub Enclosure, (20) 4" Wire Ties',  
    'refurbished',  
    'New Kicker BT2 41IK5BT2V2 Wireless Bluetooth USB Audio System Black + Remote, Power Supply (volts, ampere): 24, 2.9, Square Reflex Subwoofer (in., mm): 6 x 6", Stereo Amp Power with DSP (watts): 50, App for customizing - KickStart, Remote Control Included, Height x Width x Depth (in, mm): 8.87" x 19" x 9.14", Frequency Response (Hz): 24-20k, +-3dB, Woofer (in., cm): 5", 1 Year Manufacturer Warranty, Item Weight: 13.85 lbs., USB Port, Compatible with: Bluetooth-enabled devices, Aux-in, Speaker Design: 2-Way Full-Range, Bluetooth wireless streaming, Condition: Brand New!, Tweeter (in., cm): 3/4"],  
    dtype=object)
```

In [33]:

```
dt_electronics.pricescondition.value_counts()
```

Out[33]:

```
New  
4645  
new  
537  
Used  
118  
New other (see details)  
63  
Manufacturer refurbished  
43  
Seller refurbished  
15  
Refurbished  
8  
pre-owned  
3  
refurbished  
2  
New Kicker BT2 41IK5BT2V2 Wireless Bluetooth USB Audio System Black + Remote, Power Supply (volts, ampere): 24, 2.9, Square Reflex Subwoofer (in., mm): 6 x 6", Stereo Amp Power with DSP (watts): 50, App for customizing - KickStart, Remote Control Included, Height x Width x Depth (in, mm): 8.87" x 19" x 9.14", Frequency Response (Hz): 24-20k, +-3dB, Woofer (in., cm): 5", 1 Year Manufacturer Warranty, Item Weight: 13.85 lbs., USB Port, Compatible with: Bluetooth-enabled devices, Aux-in, Speaker Design: 2-Way Full-Range, Bluetooth wireless streaming, Condition: Brand New!, Tweeter (in., cm): 3/4"  
1  
5/16" Ring Terminal, 3 ft. 8 GA Black Ground Cable, 6 ft. Split Loom Tubing, Depth: 6.5" (165mm) (top) 11.2" (285mm) (bottom), Item Weight: 18.5 lbs., Frequency Response 25Hz - 500Hz, Line Output, Max Power: 1100 Watt x 1 Channel @ 2 Ohm, 30 ft. Speaker Wire, Boss Illuminated Logo, Height: 12.8" (325mm), (3) Rubber Grommets, Item Weight: 2 lbs., Size 10", 20 ft. 8 GA Red Power Cable, Ported enclosure for greater output and d
```

eeper bass, 2 Ohm Stable, Class A/B, Voice Coil Size 2", Black rubber surround, Nominal Impedance 4 ohm, Rugged fiberboard construction with thick carpet covering, Warranty: 1 Year Manufacturer Warranty, MOSFET Power, Weight: 6 lbs, Width: 17.7" (450mm), Condition: Brand New!, Low Pass Crossover, List item, RMS Power: 250 Watt x 1 Channel @ 4 Ohm, Remote Bass Control Included!, 1/4" Ring Terminal, 16 ft. 18 GA Blue Turn-On Wire, Peak Power: 500 Watts, Competition High Quality Fuse Holder, Condition: BRAND NEW!, Product Condition: Brand New, RMS Power: 175 Watts, Aluminum woofer cone, THD: 0.01%, 1 Year Manufacturer Warranty, Dimensions: 10-7/16" (W) x 2-1/4" (H) x 9-1/8" (L), #10 Ring Terminal, 20 ft. High Performance Black RCA, SPL (db @ 1w/1m) 88dB, New Dual SBX101 10" 500 Watts Car Audio Subwoofer Sub + Ported Sub Enclosure, (20) 4" Wire Ties 1

Name: pricescondition, dtype: int64

In [34]:

```
pricescondition_map=dt_electronics['pricescondition'].value_counts().to_dict()
pricescondition_map
```

Out[34]:

```
{'New': 4645,
 'new': 537,
 'Used': 118,
 'New other (see details)': 63,
 'Manufacturer refurbished': 43,
 'Seller refurbished': 15,
 'Refurbished': 8,
 'pre-owned': 3,
 'refurbished': 2,
 'New Kicker BT2 41IK5BT2V2 Wireless Bluetooth USB Audio System Black + Remote, Power Supply (volts, ampere): 24, 2.9, Square Reflex Subwoofer (in., mm): 6 x 6", Stereo Amp Power with DSP (watts): 50, App for customizing - KickStart, Remote Control Included, Height x Width x Depth (in, mm): 8.87" x 19" x 9.14", Frequency Response (Hz): 24-20k, +-3dB, Woofer (in., cm): 5", 1 Year Manufacturer Warranty, Item Weight: 13.85 lbs., USB Port, Compatible with: Bluetooth-enabled devices, Aux -in, Speaker Design: 2-Way Full-Range, Bluetooth wireless streaming, Condition: Brand New!, Tweeter (in., cm): 3/4": 1,
 '5/16" Ring Terminal, 3 ft. 8 GA Black Ground Cable, 6 ft. Split Loom Tubing, Depth: 6.5" (165mm) (top) 11.2" (285mm) (bottom), Item Weight: 18.5 lbs., Frequency Response 25Hz - 500Hz, Line Output, Max Power: 1100 Watt x 1 Channel @ 2 Ohm, 30 ft. Speaker Wire, Boss Illuminated Logo, Height: 12.8" (325mm), (3) Rubber Grommets, Item Weight: 2 lbs., Size 10", 20 ft. 8 GA Red Power Cable, Ported enclosure for greater output and deeper bass, 2 Ohm Stable, Class A/B, Voice Coil Size 2", Black rubber surround, Nominal Impedance 4 ohm, Rugged fiberboard construction with thick carpet covering, Warranty: 1 Year Manufacturer Warranty, MOSFET Power, Weight: 6 lbs, Width: 17.7" (450mm), Condition: Brand New!, Low Pass Crossover, List item, RMS Power: 250 Watt x 1 Channel @ 4 Ohm, Remote Bass Control Included!, 1/4" Ring Terminal, 16 ft. 18 GA Blue Turn-On Wire, Peak Power: 500 Watts, Competition High Quality Fuse Holder, Condition: BRAND NEW!, Product Condition: Brand Ne
```

w, RMS Power: 175 Watts, Aluminum woofer cone, THD: 0.01%, 1 Year Manufacturer Warranty, Dimensions: 10-7/16" (W) x 2-1/4" (H) x 9-1/8" (L), #10 Ring Terminal, 20 ft. High Performance Black RCA, SPL (db @ 1w/1m) 88dB, New Dual SBX101 10" 500 Watts Car Audio Subwoofer Sub + Ported Sub Enclosure, (20) 4" Wire Ties': 1}

In [35]:

```
dt_electronics['pricescondition']=dt_electronics['pricescondition'].map(pricescondition_map)
dt_electronics.head()
```

Out[35]:

	pricesavailability	pricescondition	pricescurrency	pricesdateSeen	pricesisSale
0	1590	4645	USD	2017-05-10T20:00:00Z,2017-05-09T15:00:00Z	False
1	1590	4645	USD	2017-10-10T02:00:00Z,2017-08-12T03:00:00Z,2017...	False
2	1590	4645	USD	2017-10-10T19:00:00Z,2017-09-12T14:00:00Z,2017...	False
3	1590	4645	USD	2017-09-08T05:00:00Z,2017-09-18T13:00:00Z,2017...	False
4	63	4645	USD	2017-12-05T13:00:00Z	True

For pricescurrency and pricesisSale I will apply one-hot encoding

In [36]:

```
dt_electronics.pricescurrency.unique()
```

Out[36]:

```
array(['USD', 'CAD'], dtype=object)
```

In [37]:

```
dt_electronics.pricesisSale.unique()
```

Out[37]:

```
array([False,  True])
```

In [38]:

```
Currency=pd.get_dummies(dt_electronics.pricescurrency,drop_first=True, prefix='Currency')
```

In [39]:

```
dt_electronics=pd.concat([dt_electronics,Currency],axis=1)
```

In [40]:

```
dt_electronics.drop('pricescurrency',axis=1,inplace=True)
```

In [41]:

```
pricesisSale=pd.get_dummies(dt_electronics.pricesisSale,drop_first=True, prefix='pricesisSale')
dt_electronics=pd.concat([dt_electronics,pricesisSale],axis=1)
dt_electronics.drop('pricesisSale',axis=1,inplace=True)
dt_electronics.head()
```

Out[41]:

	pricesavailability	pricescondition	pricesdateSeen	pricesmerchant	pricesst
0	1590	4645	2017-05-10T20:00:00Z,2017-05-09T15:00:00Z	Bestbuy.com	
1	1590	4645	2017-10-10T02:00:00Z,2017-08-12T03:00:00Z,2017...	Bestbuy.com	
2	1590	4645	2017-10-10T19:00:00Z,2017-09-12T14:00:00Z,2017...	Bestbuy.com	
3	1590	4645	2017-09-08T05:00:00Z,2017-09-18T13:00:00Z,2017...	Bestbuy.com	
4	63	4645	2017-12-05T13:00:00Z	bhphotovideo.com	Free Ex Ship most o'

I will calculate the count of each unique values of pricesmerchant and replace the value with the count

In [42]:

```
dt_electronics.pricesmerchant.unique()
```

Out[42]:

```
array(['Bestbuy.com', 'bhphotovideo.com', 'Video & Audio Center',
       'Creative Audio', 'mnmdiscount1', 'CompuDirect',
       'World Wide Stereo', 'tri-state-camera',
       'Coldriver20 Electronics Inc', 'wholesale_connection',
       'AMI Ventures Inc', 'PhotoSavings', 'EZ Tech',
       'West Coast Limited', 'Walmart.com', 'OJCommerce', 'Beach Camera',
       'Wholesale Connection', 'n.n3288', 'gabrians', 'ElectroCell',
       'legendmicro-wholesale', 'Walts TV', 'Car Audio Closeout',
       'eLGeo Electronics', 'Electronic Express', 'brandsmartusa',
       'ewholesale2011', 'evergreenpoint', 'onlycellphone',
       'hypermicrosystems', 'musicalsupplydirect', 'ericy23',
       'buydig',
       'videoandaudiocenter', 'gear4less', 'DealClock', 'wwstereo',
       'Focus Camera', 'dealonline18', 'cable-n-chargers',
       'electronicsliquidators', 'tfldealz', 'SaveCentral, LLC',
       'Clover Hill', 'waltstv', 'electricdiamond',
       'Alex Magic Electronics', 'Sound of Tri-State', 'aapltree',
       'DiamondBuy', 'Top Deals', '17th St Photo Supply Inc',
       'mike_gamesnmore', 'Marine Discount Center, LLC',
       'Technology Galaxy', 'Best Buy', 'My Goods', 'Beach Audio Inc',
       'thepixelhub', 'Sam Ash Music Direct', 'teampsales',
       'cybervalue',
       'webtron-x.com', 'bigvalueinc', 'multitechinc', 'lwin',
       'Double Deals', 'vipoutlet', 'BuyNESP.com', 'Growkart',
       'dwi-international-8', 'BuyVPC', 'stay_at_home_dads',
       '19th Hole Deals', 'Electronics Expo (Authorized Dealer)',
       'Wayfair', 'digiherd', 'Datavision Computer Video', 'K&M',
       'Real Basics', 'sundown.one', 'accesscables', 'OutletPC',
       'pmiauctions', 'dwgreatdeals', 'Circuit City - Walmart.com',
       'doitjust4fun', 'Good Guys Electronics', 'Tech2Date',
       'newopenbox',
       'Hyper Microsystems, Inc.', 'Incredible Cellular LLC',
       'Wayfair - Bestbuy.com', 'OneStopShop', 'OneCall', 'tbdeals',
       'gadget-town', 'bestserviceguys', 'goldstar_tech', 'JB Tool Sales',
       'platinummicro', 'SharpPrices', 'Adorama Camera',
       'GRAMOPHONE (Authorized Dealer)', 'antonline', 'photov
```

ideo4less',
 'AUDIO WATT STORE', 'bargainbrokers', 'jbtoolsales02',
 'electronic_express', '305mia65', 'weoutfit', 'ELAVA L
LC',
 'www-sonicelectronix-com', 'closeoutmartdeals', 'Dell.
com',
 'Soltech US Corp', 'CosmeticMall', 'raghav_205', 'mpe0
07',
 'wirelessalliance', 'redtagcamera', 'best_buy', 'techh
ouse520',
 'kays_place_too', 'Focus Camera - Walmart.com', 'blin
q',
 'ravencat3', 'TEKONLINE', 'Malelo and Company', 'Unbe
atableSale',
 'autosounds_warehouse', 'HappEshopper', 'Hot Deals 4 L
ess?',
 'LC Bargain Sales', 'refurbforless', 'E-Warehouse-Deal
s',
 'geekdeal', 'zoneusa', 'outletmp3', 'gmatexpert',
 'Unbeatable Deals', 'mamastore_co', 'echo-and-optics',
 'Monoprice Inc', 'Tech For Less Inc', 'Aiwa Inc.', 'pl
ayalegend',
 'singledadsunited', 'hsum8885', 'DemProductSales', 'np
s',
 'All in Retail', 'BuyDirect & Save', 'Adorama', 'HPP E
nterprises',
 '"Samy's Camera", "sando_distributions", "Global Audio
Video LLC",
 'lucky-31067', 'The Pixel Hub', 'Newegg', 'DLR Hunte
r',
 'proaudio654', 'Wayfair - Walmart.com', 'Olivabel',
 'Pro AV Dealer', 'new-technologies', 'Made Technolog
y', 'gpscity',
 'buynmdeals', 'kttrading14', 'Santa Monica Wireless',
 'Audio Visual Hub', 'pcRUSH - Walmart.com', 'bluedealz
-com',
 'MEMORY n MORE', 'hidedeflifestyle1', 'District Camera',
 'd.a.n.t_tech', 'Shopemco', '6Ave Express LLC', 'ipade
lectronics',
 'Fab4Electronics', 'listenup', 'istrawberrize', 'aSavi
ngs',
 'Acumen Disc', 'pro_car_audio_consulting', 'Googol Sho
p',
 'antonline.com', 'Audio Advice', 'datu503', 'outletp
c',
 'a1Components', 'DigitalKynect', 'robertscamera', 'Rac
kGo',
 'soundoftristate', 'poppy9780', 'Nexit, Inc.',
 'True Modern Electronics', 'amc0052010', 'extremenetwo
rk',
 'compuccell.tech', 'BuilderDepot, Inc.', 'OliveElect, L
LC',
 'Music123 - Walmart.com', 'Platinum Micro', 'iphoneguy
19',
 'AV Toy Store', 'Zeek and Zellies', 'ryajoh7', 'Audios
avings',
 'truemodernelectronicsusa', 'HONESTDEALS', 'djwoodie',
 'Music123',
 'jbcuttlefish', 'chribrackma-0', 'bornewinner', 'umpqu
ablia'

awtue ,
 'raffaffle', 'ejuky', 'cashinwhi', '4surpluscity', 'te
chlstore',
 'autosounds4less', 'jazzman2275', 'rideausales',
 '1 Stop Electronics Center, Inc.', 'dieg5651_8xrznj2z
i',
 'chrisdamron148', 'Always In Touch', 'beachaudio', 'tu
nnyllc',
 'PT Etech', '6ave', 'efesales', 'shopcelldeals', 'Hook
edOnTronics',
 'eCopUSA', 'iTech INTL', 'BidDeal', 'Lightning Deals',
 'ONE STOP SHOP', 'Give 5 To Cancer, Inc', 'Mian Brothe
rs Inc',
 'Tri State Camera', 'VMInnovations', 'gearbrokers', 'm
ygoods',
 'Cameta Camera', 'HiFiSoundConnection', 'Xcess Limite
d', 'axxii',
 'aimusa', 'zal-digital', 'topusedcomputers', 'g4mmlj',
'Yogi Comp',
 'a4c', 'felixg7488', 'Audiolab Stereo & Video Center',
'lite_corp',
 'Electronics101deals', 'cynethmcneil_0', 'candys-outle
t-store',
 'The Wires Zone', 'IQ Home Entertainment', 'Audiovideo
direct',
 'BlueDealz', 'ditech_store', 'cttax', 'theimagingwor
ld',
 'Nationwide Distributor', 'DigitalShopper', 'great_reta
ilers',
 'daniellcreighto_5', 'ami_lamra', 'Shoplet', 'iban_tec
h-8',
 'ovvapro', 'kourke', 'ezcellerz', 'serversourcetechnol
ogy',
 'autoplicity.com', 'Hotelectronics', 'jamhuritech', 'o
necall',
 'salherman', 'business-supplies-2009', 'universe-deal
s',
 'livinitandlovinitretail', 'claimthis', 'aad_77', 'gvi
nu',
 'pricebreak-deals', 'ProMax Commerce', 'balkan_sale',
'saving2you',
 'MyOfficeInnovations', 'redcloudtech', 'SCONTROL', 'i
electrica',
 'timespioneer2016', 'Simple Cell Inc', 'getdealonlin
e', 'atoz1239',
 'Gigasonic', 'Quality Photo', 'dealsfellow',
 'Rocky Mountain Textbooks', 'gsalabama', 'TELeasy',
 'sunshinemworld8689', 'turbopm', 'tanyastreasuries',
 'vminnovations', 'Digital Media Outlet', 'overstock557
7',
 'eGoodies', 'volaudio', 'proavdealer', 'Concept Electr
onics',
 'globalavllic', 'Spy Tec', 'department75', 'blueboxsale
s',
 'uShopMall', 'AngelSeller', '2011dailydeals', 'SaveTro
nix',
 'eternalphoto', 'mikromarc', 'ands3ts', 'extremeelectr
onix',
 'Overstock', 'pcRUSH', 'barterboxusa', 'gamegear14',
'Neobits',
 'Zazzle4573', 'Bizz12vol1', 'VTPORTFEST', 'iSavo - Walmar

zalando', 'muyizvole', 'v11001111', 't0ave', 'walmat
t.com',
 'dependix', 'Hippo-Deals', 'sound_pros', 'summerishere
-0',
 'TigerDirect - Walmart.com', 'tedselectronics1',
 'surplusstockpile', 'cyberoverstock-com', 'pillsboy300
0',
 'dinsdale-usa', 'Amazing Cross Cloud Deals', 'XOTIC P
C',
 'safeandsoundhq', 'California Electronics',
 'www-bestelectronicsoutlet-com', 'sandbid-com', 'theke
ykey',
 'AAAAA UNIVERSE', 'onlybargains2011', 'thrillingaudio',
 'CarAudioDistributors', 'Spddys LLC', 'ETM Online Deal
s, LLC',
 'Aztek Computers', 'give5tocancer', 'Alwayz On Sale',
 'acumendiscs', 'Toolboxsupply', 'guru_5', 'sqellc',
 'mainly-electronics', 'yogi-comp', 'discountparts99',
'bigdeals',
 'aloahasgt', 'Zoro', 'Amazing Deals Online', 'Lastofthe
loot',
 'Buy.com - Bestbuy.com', 'Hisense USA Corporation',
 'Computer Brain', 'Runtechmedia', 'kcla1026', '702-mot
oring',
 'Onlinecarstereo_Store', 'happychicken-com', 'JBMO Sal
es',
 'e-shopsave', 'john-6807', 'digital-solomon', 'Always.
Deals',
 'loothive', 'alex_the_terrible', 'secondipity', 'PIXIB
YTES INC',
 'superdealunlimited', 'satman56', 'grliquidators16',
 'Rover Store, Inc.', 'Mobile Top Deal', 'Amazing Savin
gs',
 'Wholesale Photo Supply Inc.', 'focuscamera', 'rabera',
 'cgelectronics16', 'pano*rama', 'Elem Technologies',
 'anjcloselouts', 'hokie_vt', 'onsale777', 'HiDEF Lifest
yle',
 'amazingtexasdeals', 'Venu', 'tierlxswccx', 'Zoe Syste
ms',
 'Good Luck Bargains', 'computeroverhauls', 'hollitroni
c',
 'irenzolio534', 'georgec4459', 'iSave', 'isellstuff8
6',
 'jakejake25', 'MegaWatts', 'Department75', 'prepre81',
 'Deals Unlimited', 'Digitaldjgear', 'hotdeals4u', 'ita
tlantic',
 'charlatan19', 'TEKENVY - Walmart.com', 'nitzafriedma
0',
 'your-best-store', 'CopyThat!', 'radha_stores', 'ltwri
ght2006',
 'maxxa1', 'northwest-outdoor', 'directnetshopdeal', 'D
eepDiscount',
 'smart_alliance', 'Widgetree', 'Moda Bella', 'discount
sjungle',
 'Maxi-Aids', 'Product Source', 'jonrod_b49gih', 'feef
f',
 'hanabeauty10', 'book2301', 'PPG Consulting, Inc',
 'ZeeMo Distribution', 'shinjuku-donquixote', 'shopdivv
y',
 'hiction', 'Tunno T.L.C', 'MemoryC Inc', 'fantoller'

```
selection , funny咄咄 , memory , etc. , lancelet ,
'cokekid88',
    'wirelessmoo', 'clocksntoys', 'pasbuy', 'qsccomputer',
    'umati_housewtp', 'brake2buylc', 'Hot Deals Electroni
cs',
    'sanheshun', 'Ritz Camera', 'trd_digital', 'alpinespri
ngers99',
    'MA Trade LLC', 'Circuit City', 'meiyaps', 'Triplenet
Pricing',
    'Deals Daily', 'Direct Home Supplies', 'GeekZilla',
    'turtlebeach_inc', '47st. Photo', 'recommercedallas',
    'race_choice', 'broomstick6', 'electric-flight-sales',
    'WeaKnees DVR Superstore', 'luckystore', 'omarafat0
9',
    'bidallies', 'Sabrent', 'ListenUp', 'Alto Music', 'acc
essoryhut2',
    'premium-electronics', 'videojam-half', 'hahiscool',
'Azano Inc',
    'aftermathoutdoor336', 'Electronics Plus', 'shipfast2d
ay',
    'DiscountCentralOnline', 'techronics-store',
    'underground_bargains', 'Advanced Gadget Inc.',
    'evolving-technologies', 'less.4.more', 'apc_sales',
    'accessorynet', 'offers.store', 'Echo and Optics', 'jp
kovacic',
    'elektech.1', 'jama2363', 'WAL-MART COM', '1topwire',
    'afdsolutions', 'us2014_bacho', 'mobileshark',
    'United Liquidators', 'digjungle', 'vcc113',
    'Urban Armor Gear Official', 'dr.pepper357', 'DVCOM LL
C',
    'jacquehuey817', 'Panda Rover', 'Roberts LP', 'mcmec
h',
    'Online King', 'bestelectronics1982', 'netroniques',
    'ProElectronics Distributing Inc.', 'kevin3sty', 'Brow
nBearTech',
    'TechMemoryStore', 'ritzcameras', 'newegg', 'altomusi
c',
    'Satellite Radio Superstore', 'daniedito', 'Firemall L
LC',
    'big_deals_4_you_2', 'storesatisfactory', 'Top Rated S
eller',
    'blucoil', 'SDC-Photo', 'Bluehaze', 'nzxtinc', 'abood.
av',
    'buy.direct', 'josh-baskin-company', 'sobeonline1',
    'Beach Audio Inc - Walmart.com', 'torchehunte0', 'tech
miaonline',
    'exclusivecharms', '428ypz', 'Mwave', 'tkservices',
    'liquid.assets', 'TDPEP', 'frontrowelectronics',
    'SSE Photo & Video', 'TopRatedSeller', 'VIP OUTLET',
'BulkBuys',
    'moviegoddess85', 'tuminatr', 'ryanblue11', 'bensdealz
7',
    'ilikeike1', 'daylighttechnologies', 'nationwidedistri
butors',
    'electr0nicsrus', 'nothingbutsavings', 'World Class Sa
les',
    'Sports&Gadgets', 'einfinityshop4', 'Murloc Express',
    'Savings Center', 'Buywireless'], dtype=object)
```

In [43]:

```
dt_electronics.pricesmerchant.value_counts()
```

Out[43]:

Bestbuy.com	2109
bhphotovideo.com	1119
Walmart.com	512
Beach Camera	153
AMI Ventures Inc	46
Focus Camera	39
buydig	36
gear4less	33
echo-and-optics	32
Beach Audio Inc	29
Best Buy	26
UnbeatableSale	26
World Wide Stereo	24
Car Audio Closeout	21
BuyVPC	21
Electronics Expo (Authorized Dealer)	20
DealClock	19
Video & Audio Center	17
electronic_express	16
Electronic Express	13
antonline.com	12
OneCall	12
The Pixel Hub	11
OutletPC	11
Newegg	10
outletmp3	10
Dell.com	10
brandsmartusa	9
DemProductSales	9
Creative Audio	9
Sound of Tri-State	9
Wholesale Connection	9
Clover Hill	9
Growkart	9
wwstereo	9
Technology Galaxy	9
My Goods	8
eLGeo Electronics	8
Hyper Microsystems, Inc.	8
HPP Enterprises	8
gadget-town	8
onecall	7
best_buy	7
Sam Ash Music Direct	7
Tech For Less Inc	7
wholesale_connection	6
Made Technology	6
Wayfair - Walmart.com	6
extremenetwork	6
Audiolab Stereo & Video Center	6
DiamondBuy	6
electronicsliquidators	5
vipoutlet	5
tbdeals	5
ElectroCell	5
aimusa	5
refurbforless	5

SELLER	REVENUE
Walts TV	5
Audiosavings	5
ielectrica	5
sundown.one	5
pcRUSH - Walmart.com	5
goldstar_tech	5
Datavision Computer Video	5
HappEshopper	5
Marine Discount Center, LLC	5
AV Toy Store	5
www-sonicelectronix-com	5
focuscamera	5
beachaudio	5
uShopMall	5
autosounds_warehouse	5
a1Components	4
SaveTronix	4
HookedOnTronics	4
17th St Photo Supply Inc	4
6Ave Express LLC	4
antonline	4
Googol Shop	4
evergreenpoint	4
redtagcamera	4
Cameta Camera	4
Monoprice Inc	4
ProMax Commerce	4
autosounds4less	4
newopenbox	4
Hippo-Deals	4
SaveCentral, LLC	4
Santa Monica Wireless	4
Adorama Camera	4
Hotelectronics	4
HONESTDEALS	4
Mian Brothers Inc	4
TEKENVY - Walmart.com	4
Samy's Camera	3
gearbrokers	3
ETM Online Deals, LLC	3
Audiovideodirect	3
Hisense USA Corporation	3
Aztek Computers	3
Alex Magic Electronics	3
Tech2Date	3
PT Etech	3
Nationwide Distributor	3
extremeelectronix	3
VIPOUTLET	3
photovideo4less	3
BidDeal	3
afdsolutions	3
n.n3288	3
listenup	3
GRAMOPHONE (Authorized Dealer)	3
Wayfair	3
dwgreatdeals	3
videoandaudiocenter	3
djwoodie	3
nonn,9720	2

FFFFYOOO	~
Nexit, Inc.	3
Pro AV Dealer	3
OJCommerce	3
Circuit City - Walmart.com	3
True Modern Electronics	3
DLR Hunter	3
6ave	3
fanteller	3
Computer Brain	3
Adorama	3
Rocky Mountain Textbooks	3
CopyThat!	3
SCONTROLC	3
iphoneguy19	3
hidedeflifestyle1	3
Simple Cell Inc	3
tfldealz	3
Rover Store, Inc.	3
Always.Deals	3
ravencat3	2
HiDEF Lifestyle	2
ONE STOP SHOP	2
ELAVA LLC	2
gamegear14	2
DeepDiscount	2
DigitalKynect	2
E-Warehouse-Deals	2
lite_corp	2
The Wires Zone	2
eCopUSA	2
Onlinecarstereo_Store	2
sound_pros	2
hokie_vt	2
outletpc	2
thepixelhub	2
kttrading14	2
electr0nicsrus	2
dealonline18	2
big_deals_4_you_2	2
Shopemco	2
MA Trade LLC	2
Coldriver20 Electronics Inc	2
Olivabel	2
timespioneer2016	2
gpscity	2
Digital Media Outlet	2
bigvalueinc	2
iSave	2
DiscountCentralOnline	2
barterboxusa	2
platinummicro	2
TigerDirect - Walmart.com	2
Music123	2
tunnyllc	2
TEKONLINE	2
globalavllc	2
Hot Deals 4 Less?	2
Audio Advice	2
AUDIO WATT STORE	2
CarAudioDistributors	2

Category	Count
book2301	2
superdealunlimited	2
iSave - Walmart.com	2
compucell.tech	2
jamhuritech	2
atoz1239	2
altomusic	2
Deals Unlimited	2
pcRUSH	2
guru_5	2
Mwave	2
bling	2
Always In Touch	2
Zoro	2
XOTIC PC	2
OliveElect, LLC	2
Tri State Camera	2
JB Tool Sales	2
serversourcetechnology	2
District Camera	2
pmiauctions	2
Overstock	2
qsccomputer	2
dwi-international-8	2
tri-state-camera	2
Amazing Cross Cloud Deals	2
mnmdiscount1	2
ericy23	2
great_retailers	2
aapltree	2
GeekZilla	2
volaudio	2
Good Luck Bargains	2
digiherd	2
Good Guys Electronics	2
ipadelectronics	2
Acumen Disc	2
zara4573	2
raffaffle	2
VMIInnovations	2
PPG Consulting, Inc	2
EZ Tech	2
universe-deals	2
Moda Bella	2
Unbeatable Deals	2
sando_distributions	2
raghav_205	2
mygoods	2
West Coast Limited	2
MemoryC, Inc.	2
DigitalShopper	2
pano*rama	2
Global Audio Video LLC	2
OneStopShop	1
ovvapro	1
buynmdeals	1
shopdivvy	1
pro_car_audio_consulting	1
shipfast2day	1
afocaloc	1

handles	+
bargainbrokers	1
gsalabama	1
jama2363	1
waltstv	1
chrismron148	1
umpquablue	1
702-motoring	1
Quality Photo	1
luckeyystore	1
autoplacity.com	1
northwest-outdoor	1
SSE Photo & Video	1
tech1store	1
give5tocancer	1
Ritz Camera	1
recommercedallas	1
jbtoolsales02	1
Toolboxsupply	1
digjungle	1
cyberoverstock-com	1
Top Rated Seller	1
musicalsuppliydirect	1
digital-solomon	1
safeandsoundhq	1
nitzafriedma_0	1
singledadsunited	1
CosmeticMall	1
smart_alliance	1
ZeeMo Distribution	1
Buy.com - Bestbuy.com	1
nps	1
truemodernelectronicsusa	1
tkservices	1
Direct Home Supplies	1
department75	1
geekdeal	1
thekeykey	1
Buywireless	1
ltwright2006	1
less.4.more	1
amazingtexasdeals	1
hahiscool	1
satman56	1
accesscables	1
offers.store	1
proaudio654	1
Wayfair - Bestbuy.com	1
kourke	1
MyOfficeInnovations	1
us2014_bacho	1
aSavings	1
surplusstockpile	1
newegg	1
alohasgt	1
g4mmlj	1
TechMemoryStore	1
Double Deals	1
teampsales	1
ewholesale2011	1
hromantic6	1

pricebreak-deals	1
Xcess Limited	1
47st. Photo	1
ryajoh7	1
Sabrent	1
WAL-MART COM	1
techronics-store	1
AngelSeller	1
mainly-electronics	1
jacquehuey817	1
Savings Center	1
MegaWatts	1
Aiwa Inc.	1
nothingbutsavings	1
sunshineworld8689	1
josh-baskin-company	1
livinitandlovinitretail	1
vminnovations	1
Maxi-Aids	1
tuminatr	1
brake2buylc	1
Sports&Gadgets	1
e-shopsave	1
daniellcreighto_5	1
ilikeike1	1
jazzman2275	1
techmiaonline	1
Panda Rover	1
BuyDirect & Save	1
daniedito	1
Real Basics	1
Concept Electronics	1
accessoryhut2	1
bluedealz-com	1
isellstuff86	1
Focus Camera - Walmart.com	1
ryanblue11	1
your-best-store	1
dealsfellow	1
sobeonline1	1
balkan_sale	1
acumendiscs	1
summerishere-0	1
cable-n-chargers	1
turbopm	1
claimthis	1
ezcellerz	1
mamastore_co	1
dependix	1
Yogi Comp	1
United Liquidators	1
TDPEP	1
getdealonline	1
DVCOM LLC	1
Echo and Optics	1
Electronics101deals	1
doitjust4fun	1
bornewinner	1
Soltech US Corp	1
webtronix.com	1

weslton .com	1
MEMORY n MORE	1
Beach Audio Inc - Walmart.com	1
accessorynet	1
blueboxsales	1
California Electronics	1
Department75	1
Electronics Plus	1
irenzolio534	1
Circuit City	1
exclusivecharms	1
Roberts LP	1
wirelessalliance	1
storesatisfactory	1
Alwayz On Sale	1
Lastoftheloot	1
onlybargains2011	1
4surpluscity	1
elektech.1	1
grliquidators16	1
ProElectronics Distributing Inc.	1
premium-electronics	1
lucky-31067	1
abood.av	1
jonrod_b49gih	1
1 Stop Electronics Center, Inc.	1
clocksntoys	1
TopRatedSeller	1
directnetshopdeal	1
istrawberrie	1
torchehunte0	1
gabrians	1
428ypz	1
electricdiamond	1
secondipity	1
Zeek and Zellies	1
trd_digital	1
cgelectronics16	1
thrillingaudio	1
iTech INTL	1
Widgetree	1
305mia65	1
ands3ts	1
stay_at_home_dads	1
shinjuku-donquixote	1
ditech_store	1
Bluehaze	1
a4c	1
frontrowelectronics	1
hsum8885	1
Gigasonic	1
Triplenet Pricing	1
mike_gamesnmore	1
john-6807	1
Amazing Deals Online	1
Runtechmedia	1
dieg5651_8xrznj2zi	1
new-technologies	1
discountsjungle	1
umati_housewtp	1
enhoehuin	1

sunnestan	1
blucoil	1
BuilderDepot, Inc.	1
computeroverhauls	1
Firemall LLC	1
PIXIBYTES INC	1
jpkovacic	1
2011dailydeals	1
pillsboy3000	1
Satellite Radio Superstore	1
liquid.assets	1
bestserviceguys	1
evolving-technologies	1
hypermicrosystems	1
1topwire	1
bigdeals	1
Urban Armor Gear Official	1
omarafat09	1
gmatexpert	1
onlycellphone	1
SDC-Photo	1
BuyNESP.com	1
netroniques	1
Fab4Electronics	1
Spddys LLC	1
LC Bargain Sales	1
IQ Home Entertainment	1
sqllc	1
Platinum Micro	1
Mobile Top Deal	1
dinsdale-usa	1
rideausales	1
kevin3sty	1
jakejake25	1
ritzcameras	1
cynethmcneil_0	1
d.a.n.t_tech	1
aad_77	1
cokekid88	1
Shoplet	1
prepre81	1
aftermathoutdoor336	1
loothive	1
Wholesale Photo Supply Inc.	1
jbcuttlefish	1
candys-outlet-store	1
electric-flight-sales	1
pasbuy	1
soundoftristate	1
PhotoSavings	1
chribrackma-0	1
19th Hole Deals	1
HiFiSoundConnection	1
RackGo	1
yogi-comp	1
cttax	1
tanyastreasuries	1
CompuDirect	1
VIP OUTLET	1
ejuky	1
+un+loboach inc	1

Customer	+
BulkBuys	1
K&M	1
tedselectronics1	1
maxxa1	1
bidallies	1
Digitaldjgear	1
amc0052010	1
Neobits	1
race_choice	1
eGoodies	1
Hot Deals Electronics	1
All in Retail	1
mobileshark	1
zal-digital	1
nzxtinc	1
lwin	1
Amazing Savings	1
buy.direct	1
sandbid-com	1
felixg7488	1
eternalphoto	1
nationwidedistributors	1
Malelo and Company	1
daylighttechnologies	1
topusedcomputers	1
moviegoddess85	1
anjcloseouts	1
Incredible Cellular LLC	1
Deals Daily	1
kays_place_too	1
apc_sales	1
mikromarc	1
SharpPrices	1
ami_lamra	1
Give 5 To Cancer, Inc	1
TELeasy	1
einfinityshop4	1
Online King	1
World Class Sales	1
saving2you	1
radha_stores	1
meiyaps	1
Azano Inc	1
wirelessmoo	1
gvinu	1
underground_bargains	1
Alto Music	1
Murloc Express	1
biction	1
Lightning Deals	1
Advanced Gadget Inc.	1
kcla1026	1
www-bestelectronicsoutlet-com	1
AAAA UNIVERSE	1
techhouse520	1
JBMO Sales	1
Product Source	1
shopcelldeals	1
theimagingworld	1
multitechinc	1

```
+-----+
legendmicro-wholesale          1
bestelectronics1982            1
videojam-half                  1
Zoe Systems                     1
mpe007                          1
iban_tech-8                    1
Audio Visual Hub               1
georgec4459                     1
hanabeauty10                   1
alex_the_terrible              1
charlatan19                     1
dr.pepper357                   1
weoutfit                        1
Venu                            1
rabera                          1
feeoff                          1
discountparts99                 1
overstock5577                  1
onsale777                       1
hotdeals4u                      1
axxii                           1
alpinespringers99               1
BlueDealz                       1
Top Deals                        1
hollitronic                     1
salherman                        1
BrownBearTech                   1
playalegend                     1
happychicken-com                1
zoneusa                          1
mcmech                           1
proavdealer                     1
vcc113                          1
closeoutmartdeals              1
buy12volt                       1
bensdealz7                      1
redcloudtech                    1
ListenUp                         1
itatlantic                      1
cybervalue                      1
business-supplies-2009          1
Elem Technologies                1
Spy Tec                          1
robertscamera                   1
cashinwhi                       1
Music123 - Walmart.com         1
Tunny LLC                        1
datu503                          1
WeaKnees DVR Superstore        1
tier1lxswccx                     1
Name: pricesmerchant, dtype: int64
```

In [44]:

```
pricesmerchant_map=dt_electronics['pricesmerchant'].value_counts().to_dict()
()
pricesmerchant_map
```

Out[44]:

```
{'Bestbuy.com': 2109,
'bphphotovideo.com': 1119,
'Walmart.com': 512,
'Beach Camera': 153,
'AMI Ventures Inc': 46,
'Focus Camera': 39,
'buydig': 36,
'gear4less': 33,
'echo-and-optics': 32,
'Beach Audio Inc': 29,
'Best Buy': 26,
'Unbeatablesale': 26,
'World Wide Stereo': 24,
'Car Audio Closeout': 21,
'BuyVPC': 21,
'Electronics Expo (Authorized Dealer)': 20,
'DealClock': 19,
'Video & Audio Center': 17,
'electronic_express': 16,
'Electronic Express': 13,
'antonline.com': 12,
'OneCall': 12,
'The Pixel Hub': 11,
'OutletPC': 11,
'Newegg': 10,
'outletmp3': 10,
'Dell.com': 10,
'brandsmartusa': 9,
'DemProductSales': 9,
'Creative Audio': 9,
'Sound of Tri-State': 9,
'Wholesale Connection': 9,
'Clover Hill': 9,
'Growkart': 9,
'wwstereo': 9,
'Technology Galaxy': 9,
'My Goods': 8,
'eLGeo Electronics': 8,
'Hyper Microsystems, Inc.': 8,
'HPP Enterprises': 8,
'gadget-town': 8,
'onecall': 7,
'best_buy': 7,
'Sam Ash Music Direct': 7,
'Tech For Less Inc': 7,
'wholesale_connection': 6,
'Made Technology': 6,
'Wayfair - Walmart.com': 6,
'extremenetwork': 6,
'Audiolab Stereo & Video Center': 6,
'DiamondBuy': 6,
'electronicsliquidators': 5,
'vipoutlet': 5,
'tbdeals': 5,
'ElectroCell': 5,
'aimusa': 5,
'refurbforless': 5,
'Walts TV': 5,
'Audiosavings': 5,
'i-electronics': 5}
```

'electrocity' : 5,
'sundown.one': 5,
'pcRUSH - Walmart.com': 5,
'goldstar_tech': 5,
'Datavision Computer Video': 5,
'HappEshopper': 5,
'Marine Discount Center, LLC': 5,
'AV Toy Store': 5,
'www-sonicelectronix-com': 5,
'focuscamera': 5,
'beachaudio': 5,
'uShopMall': 5,
'autosounds_warehouse': 5,
'a1Components': 4,
'SaveTronix': 4,
'HookedOnTronics': 4,
'17th St Photo Supply Inc': 4,
'6Ave Express LLC': 4,
'antonline': 4,
'Googol Shop': 4,
'evergreenpoint': 4,
'redtagcamera': 4,
'Cameta Camera': 4,
'Monoprice Inc': 4,
'ProMax Commerce': 4,
'autosounds4less': 4,
'newopenbox': 4,
'Hippo-Deals': 4,
'SaveCentral, LLC': 4,
'Santa Monica Wireless': 4,
'Adorama Camera': 4,
'Hotelectronics': 4,
'HONESTDEALS': 4,
'Mian Brothers Inc': 4,
'TEKENVY - Walmart.com': 4,
"Samy's Camera": 3,
'gearbrokers': 3,
'ETM Online Deals, LLC': 3,
'Audiovideodirect': 3,
'Hisense USA Corporation': 3,
'Aztek Computers': 3,
'Alex Magic Electronics': 3,
'Tech2Date': 3,
'PT Etech': 3,
'Nationwide Distributor': 3,
'extremeelectronix': 3,
'VIPOUTLET': 3,
'photovideo4less': 3,
'BidDeal': 3,
'afdsolutions': 3,
'n.n3288': 3,
'listenup': 3,
'GRAMOPHONE (Authorized Dealer)': 3,
'Wayfair': 3,
'dwgreatdeals': 3,
'videoandaudiocenter': 3,
'djwoodie': 3,
'poppy9780': 3,
'Nexit, Inc.': 3,
'Pro AV Dealer': 3,
'O.TCommerce': 3

'~~Computer~~ . . ,
'Circuit City - Walmart.com': 3,
'True Modern Electronics': 3,
'DLR Hunter': 3,
'6ave': 3,
'fanteller': 3,
'Computer Brain': 3,
'Adorama': 3,
'Rocky Mountain Textbooks': 3,
'CopyThat!': 3,
'SCONTROLC': 3,
'iphoneguy19': 3,
'hideflifestyle1': 3,
'Simple Cell Inc': 3,
'tfldealz': 3,
'Rover Store, Inc.': 3,
'Always.Deals': 3,
'ravencat3': 2,
'HiDEF Lifestyle': 2,
'ONE STOP SHOP': 2,
'ELAVA LLC': 2,
'gamegear14': 2,
'DeepDiscount': 2,
'DigitalKynect': 2,
'E-Warehouse-Deals': 2,
'lite_corp': 2,
'The Wires Zone': 2,
'eCopUSA': 2,
'Onlinecarstereo_Store': 2,
'sound_pros': 2,
'hokie_vt': 2,
'outletpc': 2,
'thepixelhub': 2,
'kttrading14': 2,
'electr0nicssrus': 2,
'dealonline18': 2,
'big_deals_4_you_2': 2,
'Shopemco': 2,
'MA Trade LLC': 2,
'Coldriver20 Electronics Inc': 2,
'Olivabel': 2,
'timespioneer2016': 2,
'gpscity': 2,
'Digital Media Outlet': 2,
'bigvalueinc': 2,
'iSave': 2,
'DiscountCentralOnline': 2,
'barterboxusa': 2,
'platinummicro': 2,
'TigerDirect - Walmart.com': 2,
'Music123': 2,
'tunnyllc': 2,
'TEKONLINE': 2,
'globalavllc': 2,
'Hot Deals 4 Less?': 2,
'Audio Advice': 2,
'AUDIO WATT STORE': 2,
'CarAudioDistributors': 2,
'book2301': 2,
'superdealunlimited': 2,
'iSave - Walmart.com' : ?

+wave_watmax.com . 4,
'compucell.tech': 2,
'jamhuritech': 2,
'atoz1239': 2,
'altomusic': 2,
'Deals Unlimited': 2,
'pcRUSH': 2,
'guru_5': 2,
'Mwave': 2,
'bling': 2,
'Always In Touch': 2,
'Zoro': 2,
'XOTIC PC': 2,
'OliveElect, LLC': 2,
'Tri State Camera': 2,
'JB Tool Sales': 2,
'serversourcetechnology': 2,
'District Camera': 2,
'pmiauctions': 2,
'Overstock': 2,
'qsccomputer': 2,
'dwi-international-8': 2,
'tri-state-camera': 2,
'Amazing Cross Cloud Deals': 2,
'mnmdiscount1': 2,
'ericy23': 2,
'great_retailers': 2,
'aapltree': 2,
'GeekZilla': 2,
'volaudio': 2,
'Good Luck Bargains': 2,
'digiherd': 2,
'Good Guys Electronics': 2,
'ipadelectronics': 2,
'Acumen Disc': 2,
'zara4573': 2,
'raffaffle': 2,
'VMInnovations': 2,
'PPG Consulting, Inc': 2,
'EZ Tech': 2,
'universe-deals': 2,
'Moda Bella': 2,
'Unbeatable Deals': 2,
'sando_distributions': 2,
'raghav_205': 2,
'mygoods': 2,
'West Coast Limited': 2,
'MemoryC, Inc.': 2,
'DigitalShopper': 2,
'pano*rama': 2,
'Global Audio Video LLC': 2,
'OneStopShop': 1,
'ovvapro': 1,
'buynmdeals': 1,
'shopdivvy': 1,
'pro_car_audio_consulting': 1,
'shipfast2day': 1,
'efesales': 1,
'bargainbrokers': 1,
'gsalabama': 1,
'49mb2263': 1

'jumazoso': 1,
'waltstv': 1,
'chrisdamron148': 1,
'umpquablue': 1,
'702-motorring': 1,
'Quality Photo': 1,
'luckystore': 1,
'autoplicity.com': 1,
'northwest-outdoor': 1,
'SSE Photo & Video': 1,
'tech1store': 1,
'give5tocancer': 1,
'Ritz Camera': 1,
'recommercedallas': 1,
'jbtoolsales02': 1,
'Toolboxsupply': 1,
'digjungle': 1,
'cyberoverstock-com': 1,
'Top Rated Seller': 1,
'musicalsupplydirect': 1,
'digital-solomon': 1,
'safeandsoundhq': 1,
'nitzafriedma_0': 1,
'singledadsunited': 1,
'CosmeticMall': 1,
'smart_alliance': 1,
'ZeeMo Distribution': 1,
'Buy.com - Bestbuy.com': 1,
'nps': 1,
'truemodernelectronicsusa': 1,
'tkservices': 1,
'Direct Home Supplies': 1,
'department75': 1,
'geekdeal': 1,
'thekeykey': 1,
'Buywireless': 1,
'ltwright2006': 1,
'less.4.more': 1,
'amazingtexasdeals': 1,
'hahiscool': 1,
'satman56': 1,
'accesscables': 1,
'offers.store': 1,
'proaudio654': 1,
'Wayfair - Bestbuy.com': 1,
'kourke': 1,
'MyOfficeInnovations': 1,
'us2014_bacho': 1,
'aSavings': 1,
'surplusstockpile': 1,
'newegg': 1,
'alohasgt': 1,
'g4mmlj': 1,
'TechMemoryStore': 1,
'Double Deals': 1,
'teampsales': 1,
'ewholesale2011': 1,
'broomstick6': 1,
'pricebreak-deals': 1,
'Xcess Limited': 1,
'A7et_Photo': 1

'ryajoh7': 1,
'Sabrent': 1,
'WAL-MART COM': 1,
'techronics-store': 1,
'AngelSeller': 1,
'mainly-electronics': 1,
'jacquehuey817': 1,
'Savings Center': 1,
'MegaWatts': 1,
'Aiwa Inc.': 1,
'nothingbut savings': 1,
'sunshineworld8689': 1,
'josh-baskin-company': 1,
'livinitandlovinitretail': 1,
'vminnovations': 1,
'Maxi-Aids': 1,
'tuminatr': 1,
'brake2buylc': 1,
'Sports&Gadgets': 1,
'e-shopsave': 1,
'daniellcreighto_5': 1,
'ilikeike1': 1,
'jazzman2275': 1,
'techmiaonline': 1,
'Panda Rover': 1,
'BuyDirect & Save': 1,
'daniedito': 1,
'Real Basics': 1,
'Concept Electronics': 1,
'accessoryhut2': 1,
'bluedealz-com': 1,
'isellstuff86': 1,
'Focus Camera - Walmart.com': 1,
'ryanblue11': 1,
'your-best-store': 1,
'dealsfellow': 1,
'sobeonline1': 1,
'balkan_sale': 1,
'acumendiscs': 1,
'summerishere-0': 1,
'cable-n-chargers': 1,
'turbopm': 1,
'claimthis': 1,
'ezcellerz': 1,
'mamastore_co': 1,
'dependix': 1,
'Yogi Comp': 1,
'United Liquidators': 1,
'TDPEP': 1,
'getdealonline': 1,
'DVCOM LLC': 1,
'Echo and Optics': 1,
'Electronics101deals': 1,
'doitjust4fun': 1,
'bornewinner': 1,
'Soltech US Corp': 1,
'webtron-x.com': 1,
'MEMORY n MORE': 1,
'Beach Audio Inc - Walmart.com': 1,
'accessorunet': 1

'accessorytree': 1,
'blueboxsales': 1,
'California Electronics': 1,
'Department75': 1,
'Electronics Plus': 1,
'irenzolio534': 1,
'Circuit City': 1,
'exclusivecharms': 1,
'Roberts LP': 1,
'wirelessalliance': 1,
'storesatisfactory': 1,
'Alwayz On Sale': 1,
'Lastoftheloot': 1,
'onlybargains2011': 1,
'4surpluscity': 1,
'elektech.1': 1,
'grliquidators16': 1,
'ProElectronics Distributing Inc.': 1,
'premium-electronics': 1,
'lucky-31067': 1,
'abood.av': 1,
'jonrod_b49gih': 1,
'1 Stop Electronics Center, Inc.': 1,
'clocksntoys': 1,
'TopRatedSeller': 1,
'directnetshopdeal': 1,
'istrawberrize': 1,
'torchehunte0': 1,
'gabrians': 1,
'428ypz': 1,
'electricdiamond': 1,
'secondipity': 1,
'Zeek and Zellies': 1,
'trd_digital': 1,
'cgelectronics16': 1,
'thrillingaudio': 1,
'iTech INTL': 1,
'Widgetree': 1,
'305mia65': 1,
'ands3ts': 1,
'stay_at_home_dads': 1,
'shinjuku-donquixote': 1,
'ditech_store': 1,
'Bluehaze': 1,
'a4c': 1,
'frontrowelectronics': 1,
'hsum8885': 1,
'Gigasonic': 1,
'Triplenet Pricing': 1,
'mike_gamesnmore': 1,
'john-6807': 1,
'Amazing Deals Online': 1,
'Runtechmedia': 1,
'dieg5651_8xrznj2zi': 1,
'new-technologies': 1,
'discountsjungle': 1,
'umati_housewtp': 1,
'sanheshun': 1,
'blucoil': 1,
'BuilderDepot, Inc.': 1,
'computeroverhaul': 1

'computeroverhauls': 1,
'Firemall LLC': 1,
'PIXIBYTES INC': 1,
'jpkovacic': 1,
'2011dailydeals': 1,
'pillsboy3000': 1,
'Satellite Radio Superstore': 1,
'liquid.assets': 1,
'bestserviceguys': 1,
'evolving-technologies': 1,
'hypermicrosystems': 1,
'1topwire': 1,
'bigdeals': 1,
'Urban Armor Gear Official': 1,
'omarafat09': 1,
'gmatexpert': 1,
'onlycellphone': 1,
'SDC-Photo': 1,
'BuyNESP.com': 1,
'netroniques': 1,
'Fab4Electronics': 1,
'Spddys LLC': 1,
'LC Bargain Sales': 1,
'IQ Home Entertainment': 1,
'sqellc': 1,
'Platinum Micro': 1,
'Mobile Top Deal': 1,
'dinsdale-usa': 1,
'rideausales': 1,
'kevin3sty': 1,
'jakejake25': 1,
'ritzcameras': 1,
'cynethmcneil_0': 1,
'd.a.n.t_tech': 1,
'aad_77': 1,
'cokekid88': 1,
'Shoplet': 1,
'prepre81': 1,
'aftermathoutdoor336': 1,
'loothive': 1,
'Wholesale Photo Supply Inc.': 1,
'jbcuttlefish': 1,
'candys-outlet-store': 1,
'electric-flight-sales': 1,
'pasbuy': 1,
'soundoftristate': 1,
'PhotoSavings': 1,
'chribrackma-0': 1,
'19th Hole Deals': 1,
'HiFiSoundConnection': 1,
'RackGo': 1,
'yogi-comp': 1,
'cttax': 1,
'tanyastreasuries': 1,
'CompuDirect': 1,
'VIP OUTLET': 1,
'ejuky': 1,
'turtlebeach_inc': 1,
'BulkBuys': 1,
'K&M': 1,
'+edselectronics1': 1

'bestelectronicsoutlet.com': 1,
'maxxal': 1,
'bidallies': 1,
'Digitaldjgear': 1,
'amc0052010': 1,
'Neobits': 1,
'race_choice': 1,
'eGoodies': 1,
'Hot Deals Electronics': 1,
'All in Retail': 1,
'mobileshark': 1,
'zal-digital': 1,
'nzxtinc': 1,
'1win': 1,
'Amazing Savings': 1,
'buy.direct': 1,
'sandbid-com': 1,
'felixg7488': 1,
'eternalphoto': 1,
'nationwideditistributors': 1,
'Malelo and Company': 1,
'daylighttechnologies': 1,
'topusedcomputers': 1,
'moviegodess85': 1,
'anjcloseouts': 1,
'Incredible Cellular LLC': 1,
'Deals Daily': 1,
'kays_place_too': 1,
'apc_sales': 1,
'mikromarc': 1,
'SharpPrices': 1,
'ami_lamra': 1,
'Give 5 To Cancer, Inc': 1,
'TELeasy': 1,
'einfinityshop4': 1,
'Online King': 1,
'World Class Sales': 1,
'saving2you': 1,
'radha_stores': 1,
'meiyaps': 1,
'Azano Inc': 1,
'wirelessmoo': 1,
'gvinu': 1,
'underground_bargains': 1,
'Alto Music': 1,
'Murloc Express': 1,
'biction': 1,
'Lightning Deals': 1,
'Advanced Gadget Inc.': 1,
'kcla1026': 1,
'www-bestelectronicsoutlet-com': 1,
'AAAA UNIVERSE': 1,
'techhouse520': 1,
'JBMO Sales': 1,
'Product Source': 1,
'shopcelldeals': 1,
'theimagingworld': 1,
'multitechinc': 1,
'legendmicro-wholesale': 1,
'bestelectronics1982': 1,
'widewhale': 1

```
vaeo_jam_mall . +,  
'Zoe Systems': 1,  
'mpe007': 1,  
'iban_tech-8': 1,  
'Audio Visual Hub': 1,  
'georgec4459': 1,  
'hanabeauty10': 1,  
'alex_the_terrible': 1,  
'charlatan19': 1,  
'dr.pepper357': 1,  
'weoutfit': 1,  
'Venu': 1,  
'rabera': 1,  
'feeoff': 1,  
'discountparts99': 1,  
'overstock5577': 1,  
'onsale777': 1,  
'hotdeals4u': 1,  
'axxii': 1,  
'alpinespringers99': 1,  
'BlueDealz': 1,  
'Top Deals': 1,  
'hollitronic': 1,  
'salherman': 1,  
'BrownBearTech': 1,  
'playalegend': 1,  
'happychicken-com': 1,  
'zoneusa': 1,  
'mcmech': 1,  
'proavdealer': 1,  
'vccl13': 1,  
'closeoutmartdeals': 1,  
'buy12volt': 1,  
'bensdealz7': 1,  
'redcloudtech': 1,  
'ListenUp': 1,  
'itatlantic': 1,  
'cybervalue': 1,  
'business-supplies-2009': 1,  
'Elem Technologies': 1,  
'Spy Tec': 1,  
'robertscamera': 1,  
'cashinwhi': 1,  
'Music123 - Walmart.com': 1,  
'Tunny LLC': 1,  
'datu503': 1,  
'WeaKnees DVR Superstore': 1,  
'tier1xswccx': 1}
```

here I will replace 'Bestbuy.com' with 2109

In [45]:

```
dt_electronics['pricesmerchant']=dt_electronics['pricesmerchant'].map(pricesmerchant_map)  
dt_electronics.head()
```

Out[45]:

	pricesavailability	pricescondition	pricesdateSeen	pricesmerchant	pricesshipping
0	1590	4645	2017-05-10T20:00:00Z,2017-05-09T15:00:00Z	2109	
1	1590	4645	2017-10-10T02:00:00Z,2017-08-12T03:00:00Z,2017...	2109	
2	1590	4645	2017-10-10T19:00:00Z,2017-09-12T14:00:00Z,2017...	2109	
3	1590	4645	2017-09-08T05:00:00Z,2017-09-18T13:00:00Z,2017...	2109	
4	63	4645	2017-12-05T13:00:00Z	1119	Free Expedited Shipping for most orders over \$49

This column indicate shipping charges of each delivery

In [46]:

```
dt_electronics.pricesshipping.unique()
```

Out[46]:

```
array([nan, 'Free Expedited Shipping for most orders over $49', 'Freight', 'Standard', 'FREE', 'Value', 'Expedited', 'Free Shipping on orders 35 and up', 'Free Expedited Shipping', 'USD 3.99', 'USD 13.32', 'Free Delivery', 'Free Shipping on orders 35 and up', 'Shipping Charges Apply', 'USD 3.39', 'USD 16.00', 'USD 12.78', 'USD 13.60', 'USD 11.30', 'Free Shipping', 'USD 6.65', 'USD 15.42', 'CAD 18.61', 'USD 4.88', 'Free Standard Shipping', 'USD 10.00', 'USD 13.61', 'USD 25.00', 'USD 150.00', 'Free Standard Shipping on Orders Over $49', 'USD 196.92', 'USD 10.10', 'USD 7.81', 'USD 24.04', 'U
```

```
SD 22.64',
    'USD 75.00', 'USD 16.45', 'USD 179.35', 'USD 27.94',
'USD 15.68',

    'USD 37.65', 'USD 18.60', 'USD 13.95', 'USD 79.42',
    'Free Shipping for this Item', 'USD 5.00', 'USD 140.0
0',
    'USD 7.25', 'USD 14.00', 'USD 35.03', 'USD 0.99', 'USD
25.76',
    'USD 6.00', 'Free Next Day Delivery (USA)', 'USD 2.9
9',
    'USD 20.00', 'USD 22.13', 'USD 60.00', 'USD 2.85', 'US
D 5.99'],
    dtype=object)
```

I will replace those values which indicate free delivery with 0

In [47]:

```
dt_electronics['pricesshipping']=dt_electronics['pricesshipping'].replace
(['Free Expedited Shipping for most orders over $49','Free Shipping on ord
ers 35 and up','Free Expedited Shipping','Free Delivery','Free Shippingon
orders 35 and up','Free Shipping','Free Standard Shipping','Free Standard
Shipping on Orders Over $49','Free Shipping for this Item','Free Next Day
Delivery (USA) ','FREE'],0)
```

In [48]:

```
dt_electronics.pricesshipping.unique()
```

Out[48]:

```
array([nan, 0, 'Freight', 'Standard', 'Value', 'Expedited',
'USD 3.99',
    'USD 13.32', 'Shipping Charges Apply', 'USD 3.39', 'US
D 16.00',
    'USD 12.78', 'USD 13.60', 'USD 11.30', 'USD 6.65', 'US
D 15.42',
    'CAD 18.61', 'USD 4.88', 'USD 10.00', 'USD 13.61', 'US
D 25.00',
    'USD 150.00', 'USD 196.92', 'USD 10.10', 'USD 7.81',
'USD 24.04',
    'USD 22.64', 'USD 75.00', 'USD 16.45', 'USD 179.35',
'USD 27.94',
    'USD 15.68', 'USD 37.65', 'USD 18.60', 'USD 13.95', 'U
SD 79.42',
    'USD 5.00', 'USD 140.00', 'USD 7.25', 'USD 14.00', 'US
D 35.03',
    'USD 0.99', 'USD 25.76', 'USD 6.00', 'USD 2.99', 'USD
20.00',
    'USD 22.13', 'USD 60.00', 'USD 2.85', 'USD 5.99'], dty
pe=object)
```

Now I will remove the USD string from each value

In [49]:

```
dt_electronics['pricesshipping']=dt_electronics['pricesshipping'].replace(['USD 3.99', 'USD 13.32', 'USD 3.39', 'USD 16.00',
    'USD 12.78', 'USD 13.60', 'USD 11.30', 'USD 6.65', 'USD 15.42',
    'CAD 18.61', 'USD 4.88', 'USD 10.00', 'USD 13.61', 'USD 25.00',
    'USD 150.00', 'USD 196.92', 'USD 10.10', 'USD 7.81', 'USD 24.04',
    'USD 22.64', 'USD 75.00', 'USD 16.45', 'USD 179.35', 'USD 27.94',
    'USD 15.68', 'USD 37.65', 'USD 18.60', 'USD 13.95', 'USD 79.42',
    'USD 5.00', 'USD 140.00', 'USD 7.25', 'USD 14.00', 'USD 35.03',
    'USD 0.99', 'USD 25.76', 'USD 6.00', 'USD 2.99', 'USD 20.00',
    'USD 22.13', 'USD 60.00', 'USD 2.85', 'USD 5.99'],[3.99,13.32,3.39,
16.00,12.78,13.60,11.30,6.65,15.42,18.61,4.88,10.00,13.61,25.00,150.00,19
6.92,10.10,7.81,24.04,22.64,75.00,16.45,179.35,27.94,15.68,37.65,18.60,13.
95,79.42,5.00,140.00,7.25,14.00,35.03,0.99,25.76,6.00,2.99,20.00,22.13,60.
00,2.85,5.99])
```

I will replace all other value with the mean

To do that i will replace all the unique string type values with NaN

In [50]:

```
dt_electronics['pricesshipping']=dt_electronics['pricesshipping'].replace(['Freight','Standard','Value','Expedited','Shipping Charges Apply'],np.nan)
```

In [51]:

```
dt_electronics.pricesshipping.unique()
```

Out[51]:

```
array([  nan,  0. ,   3.99,  13.32,   3.39,  16. ,  12.7
8,  13.6 ,
       11.3 ,   6.65,  15.42,  18.61,   4.88,  10. ,  13.6
1,  25. ,
       150. , 196.92, 10.1 ,   7.81,  24.04,  22.64,  75.
,  16.45,
       179.35, 27.94, 15.68,  37.65,  18.6 ,  13.95,  79.4
2,  5. ,
       140. ,   7.25, 14. ,  35.03,   0.99,  25.76,   6.
,  2.99,
       20. , 22.13,  60. ,   2.85,   5.99])
```

Now I will find the mean and replace the nan values

In [52]:

```
mean=dt_electronics.pricesshipping.mean()
mean=round(mean)
```

In [53]:

```
dt_electronics[['pricesshipping']] = dt_electronics[['pricesshipping']].fillna(mean, axis=1)
```

In [54]:

```
dt_electronics['pricesshipping'] = dt_electronics['pricesshipping'].astype('int64')
```

In [55]:

```
dt_electronics.pricesshipping.unique()
```

Out[55]:

```
array([ 1,  0,  3, 13, 16, 12, 11,  6, 15, 18,  4,
10, 25,
       150, 196,  7, 24, 22, 75, 179, 27, 37, 79,  5,
140, 14,
       35,   2, 20, 60], dtype=int64)
```

Now for brand column I provide Rank to each unique brand

In [56]:

```
dt_electronics.brand.unique()
```

Out[56]:

```
array(['Grace Digital', 'Lenovo', 'House of Marley', 'Sony',
'Acer', 'LG',
       'Yamaha', 'Kicker', 'Turtle Beach', 'Russound', 'Sandisk',
       'Logitech', 'Samsung', 'WD', 'Nikon', 'ASUS', 'Thermal take',
       'Joby', 'Power Acoustik', 'Sonic Alert', 'Elite Screens',
       'Klipsch', 'RODE', 'SiriusXM', 'Alpine', 'Razer', 'Motorola Home',
       'Polk Audio', 'CORSAIR', 'Pioneer', 'StarTech.com', 'JBL',
       'Jvc', 'Canon', 'Apple', 'Cerwin Vega', 'Panamax', 'Denon',
'iLive',
       'Virgin Mobile', 'Seagate', 'Kensington', 'SVS', 'Tp-Link',
       'Garmin', 'SanDisk', 'Pny', 'Polaroid', 'Isimple',
       'Universal Remote Control', 'Sylvania', 'Sharp', 'Alienware',
       'Dell', 'Hisense', 'Westone', 'Escort', 'Kanto Living',
       'SOL REPUBLIC', 'Panasonic', 'Kanto', 'Belkin', 'Peerless',
       'Zoom',
       'Sennheiser', 'DENAQ', 'Sanus', 'Bose®', 'Western Digital',
       'Sling Media', 'Sima', 'Netgear', 'V-MODA', 'Manfrotto',
       'Monster',
       'Zmodo', 'VIZIO', 'Boytone', 'Siriusxm', 'ECOXGEAR',
'Manufactory'])
```

'MARTINLOGAN',
 'XFX', 'HP', 'Olympus', 'Microsoft', 'Corsair', 'Waco
m', 'Dual',
 'ZTE', 'V-Moda', 'Pyle Pro', 'Cerwin-Vega', 'Onkyo',
'NVIDIA',
 'CLARITY-TELECOM', 'Asus', 'iSimple', 'Definitive Tech
nology',
 'Linksys', 'PyleHome', 'Kenwood', 'Braven', 'Jbl', 'MS
I',
 'Lowepro', 'G-Technology', 'Optoma', 'Toshiba', 'GEK
O',
 'Sideclick', 'Actiontec', 'ViewSonic', 'Buffalo',
 'Cobra Electronics', 'Silicondustry', 'SunBriteTV', 'TP-
Link',
 'Cooler Master', 'VINPOWER DIGITAL DIRECTSHIP', 'Midla
nd',
 'Visidec', 'Bower', 'Sabrent', 'Klipsch Xr8i In-Ear He
adphones',
 'Jabra', 'mophie', 'BIC America', 'Peak Design', 'Jens
en',
 'Marantz', 'Belkin Inc.', 'Viper', 'AudioQuest', 'Case
Logic',
 'Sound Design', 'kate spade new york', 'Buffalo Techno
logy',
 'Bose', 'GoPro', 'KICKER', 'Tivo', 'Insignia', 'RCA',
'Beats',
 'Pyle', 'Tiffen', 'Secur', 'Sol Republic', '360fly',
'Steelseries',
 'Mevo', 'Aiwa', 'Sherwood', 'MEE audio', 'Peerless-A
V', 'Ghostek',
 'Epson', 'Sigma', 'ZAGG', 'Kingston', 'MTX Audio', 'Sp
artan',
 'Speck', 'Atrend', 'TCL', 'Sengled', 'Memorex', 'Ante
c',
 'Rand McNally', 'StarTech', 'Aftershokz', 'MTX', 'Nova
tel',
 'Google', 'LifeProof', 'SANUS', 'Innovative Technolog
y', 'AOC',
 'Arris', 'PNY', 'Master Dynamic', 'Aluratek', 'PELICA
N',
 'Digipower', 'House Of Marley', 'Anker', 'TiVo', 'Visi
onTek',
 'BOYO', 'Transcend', 'Jaybird', 'Urban Armor Gear', 'N
ZXT',
 'IOGEAR', 'Audioquest', 'SpeakerCraft', 'SKB', 'I.am
+',
 'Planet Audio', 'FUGOO', 'Zubie', 'SHARKK', 'Goal Zer
o',
 'PreSonus', 'Mee Audio', 'Niles', 'Fusion Marine', 'NE
TGEAR',
 'M-Audio', 'Sunpak', 'ONLINE', 'Fitbit', 'Papago', 'So
nos', 'Evga',
 'Tenba', 'X6D', 'ECO STYLE', 'Outdoor Tech', 'Cobra',
 'DENON - HEOS', 'CybertronPC', 'Chief', 'Ultimate Ear
s', 'JVC',
 'Diamond Multimedia', 'Hoya', 'PANAMX', 'Intel', 'D-Li
nk',
 'Sdi Technologies, Inc.', 'BenQ', 'Master & Dynamic',
'Lorex',
 'Crosley', 'KEF', 'Thule', 'Motorola', 'Sonax', 'Lee
f'

```
+     'Hauppauge', 'AfterShokz', 'Incipio', 'OmniMount', 'Vi
ctrola',
     'PYLE', 'Insignia™', 'Skullcandy', 'Skybell', 'Energiz
er',
     'j5create', 'VXi', 'Retrak', 'DreamWave'], dtype=objec
t)
```

In [57]:

```
dt_electronics.brand.value_counts()
```

Out[57]:

Sony	578
Samsung	565
Apple	188
Yamaha	183
Pioneer	137
LG	132
Logitech	82
Lenovo	80
Kenwood	74
SanDisk	74
CORSAIR	73
Sennheiser	72
WD	72
Alpine	72
Corsair	67
Elite Screens	65
Razer	62
ASUS	58
Seagate	57
Onkyo	55
Panasonic	52
Lowepro	50
Netgear	50
Alienware	49
Dell	46
JBL	45
Denon	43
TP-Link	42
Power Acoustik	41
Garmin	38
Sandisk	37
Russound	36
Western Digital	36
House of Marley	35
Olympus	34
Canon	34
Definitive Technology	33
Kicker	32
V-Moda	32
Epson	31
Microsoft	28
Grace Digital	27
AudioQuest	25
ECOXGEAR	25
Actiontec	24
SOL REPUBLIC	24

DENAQ	24
Sharp	23
Joby	23
Asus	22
Siriusxm	22
Zoom	21
Acer	21
Kanto Living	21
MEE audio	21
AOC	20
Jbl	20
Nikon	20
Hisense	19
Klipsch	19
Sylvania	19
ViewSonic	18
Sanus	18
Pyle Pro	18
MartinLogan	18
360fly	18
V-MODA	18
Cooler Master	17
Steelseries	17
iSimple	16
Bose	16
MSI	16
Insignia	16
Optoma	16
Pny	15
Thermaltake	15
Bose®	15
Boytone	15
Peak Design	15
GEKO	15
Toshiba	15
mophie	14
NZXT	14
Rand McNally	14
Polk Audio	14
Bower	14
Manfrotto	13
Silicondust	13
Linksys	13
HP	13
Belkin	13
StarTech.com	13
Kanto	13
Papago	13
Isimple	13
SunBriteTV	13
Kingston	13
Belkin Inc.	12
SVS	12
VIZIO	12
Pyle	12
Marantz	12
Wacom	12
BIC America	11
Zmodo	11
Jensen	11
G-Technology	11

Company	Count
NETGEAR	11
LifeProof	11
Westone	11
RCA	10
ZTE	10
Visidec	10
Cobra Electronics	10
JvcM	10
Spartan	10
Google	10
Sima	10
MTX	10
Panamax	10
Atrend	10
Monster	9
Sherwood	9
SiriusXM	9
Mevo	9
PyleHome	9
Sigma	9
Intel	9
Peerless-AV	9
Braven	9
XFX	9
RODE	9
Arris	8
Midland	8
Tenba	8
Escort	8
ZAGG	8
Sound Design	8
Motorola	8
Niles	8
Jaybird	8
Tiffen	8
Dual	8
SHARKK	8
Sengled	8
VisionTek	7
Peerless	7
Motorola Home	7
DENON - HEOS	7
Viper	7
House Of Marley	7
GoPro	7
PYLE	7
CybertronPC	7
Cerwin Vega	7
PNY	6
Speck	6
TCL	6
Polaroid	6
PELICAN	6
Sideclick	6
CLARITY-TELECOM	6
NVIDIA	6
Anker	6
Aluratek	6
Cobra	6
CKR	6

SIRI	5
Cerwin-Vega	6
StarTech	6
MTX Audio	5
M-Audio	5
Urban Armor Gear	5
iLive	5
Sonos	5
IOGEAR	5
Tivo	5
SANUS	5
Outdoor Tech	5
Leef	5
Mee Audio	5
Buffalo Technology	5
KICKER	5
BOYO	5
Antec	5
Novatel	5
Thule	4
FUGOO	4
Aiwa	4
Audioquest	4
Hoya	4
Kensington	4
Klipsch Xr8i In-Ear Headphones	4
Universal Remote Control	4
Planet Audio	4
kate spade new york	4
Transcend	4
Evga	4
Master Dynamic	4
Turtle Beach	4
Sonic Alert	4
Retrak	4
ONLINE	4
Buffalo	4
Fitbit	4
Skybell	4
Fusion Marine	4
SpeakerCraft	4
X6D	3
Ultimate Ears	3
Sling Media	3
Insignia™	3
Master & Dynamic	3
Jabra	3
Memorex	3
Beats	3
OmniMount	3
Secur	3
TiVo	3
Virgin Mobile	3
Crosley	3
ECO STYLE	3
Tp-Link	3
Sdi Technologies, Inc.	3
AfterShokz	3
PreSonus	3
Sabrent	3
D-Tech	2

```
Digipower          2
JVC               2
Goal Zero         2
Aftershokz        2
Incipio            2
Ghostek            2
Innovative Technology 2
VINPOWER DIGITAL DIRECTSHIP 2
Case Logic          2
Chief               2
Sol Republic        2
Sunpak              2
Skullcandy          2
BenQ                2
KEF                 2
VXi                 1
Victrola            1
DreamWave           1
I.am+               1
Hauppauge           1
Lorex               1
Sonax               1
Zubie               1
Energizer            1
j5create             1
PANAMX              1
Diamond Multimedia  1
Name: brand, dtype: int64
```

I will get this rank with respect to price column

In [58]:

```
ordinal_labels=dt_electronics.groupby(['brand'])['price'].mean().sort_values().index
```

In [59]:

```
ordinal_labels
```

Out[59]:

```
Index(['Retrak', 'Sunpak', 'Secur', 'iLive', 'Energizer', 'RC
A', 'iSimple',
       'Belkin Inc.', 'DENAQ', 'Case Logic',
       ...
       'Apple', 'Samsung', 'Optoma', 'Alienware', 'Asus', 'Ma
rantz', 'MSI',
       'Intel', 'LG', 'SunBriteTV'],
      dtype='object', name='brand', length=254)
```

In [60]:

```
enumerate(ordinal_labels,0)
```

Out[60]:

```
<enumerate at 0x21d45f67c18>
```

In [61]:

```
ordinal_labels2={k:i for i,k in enumerate(ordinal_labels,0)}  
ordinal_labels2
```

Out[61]:

```
{'Retrak': 0,  
 'Sunpak': 1,  
 'Secur': 2,  
 'iLive': 3,  
 'Energizer': 4,  
 'RCA': 5,  
 'iSimple': 6,  
 'Belkin Inc.': 7,  
 'DENAQ': 8,  
 'Case Logic': 9,  
 'Leef': 10,  
 'Speck': 11,  
 'j5create': 12,  
 'GoPro': 13,  
 'Tiffen': 14,  
 'Sideclick': 15,  
 'Sonax': 16,  
 'Midland': 17,  
 'X6D': 18,  
 'Viper': 19,  
 'Anker': 20,  
 'Sabrent': 21,  
 'Sonic Alert': 22,  
 'TP-Link': 23,  
 'CLARITY-TELECOM': 24,  
 'Sdi Technologies, Inc.': 25,  
 'Insignia': 26,  
 'BOYO': 27,  
 'Kingston': 28,  
 'PNY': 29,  
 'Mee Audio': 30,  
 'Thule': 31,  
 'Aftershokz': 32,  
 'Peerless-AV': 33,  
 'Memorex': 34,  
 'Insignia™': 35,  
 'IOGEAR': 36,  
 'SKB': 37,  
 'Pyle Pro': 38,  
 'Joby': 39,  
 'Bower': 40,  
 'Ghostek': 41,  
 'Antec': 42,  
 'Boytone': 43,  
 'OmniMount': 44,  
 'AfterShokz': 45,  
 'Peerless': 46,  
 'SHARKK': 47,  
 'Digipower': 48,  
 'Victrola': 49,  
 'kate spade new york': 50.}
```

'Incipio': 51,
'PELICAN': 52,
'Isimple': 53,
'Thermaltake': 54,

'Urban Armor Gear': 55,
'Outdoor Tech': 56,
'Skullcandy': 57,
'Wacom': 58,
'GEKO': 59,
'VisionTek': 60,
'Lowepro': 61,
'Novatel': 62,
'mophie': 63,
'Turtle Beach': 64,
'LifeProof': 65,
'Zmodo': 66,
'Manfrotto': 67,
'Actiontec': 68,
'Hauppauge': 69,
'Tp-Link': 70,
'Sound Design': 71,
'Dual': 72,
'Google': 73,
'Fusion Marine': 74,
'Kensington': 75,
'ECO STYLE': 76,
'Atrend': 77,
'Hoya': 78,
'Sylvania': 79,
'Motorola': 80,
'Jvcn': 81,
'Monster': 82,
'MTX': 83,
'Siriusxm': 84,
'AOC': 85,
'Innovative Technology': 86,
'NZXT': 87,
'Aluratek': 88,
'Peak Design': 89,
'Microsoft': 90,
'PANAMX': 91,
'Diamond Multimedia': 92,
'Grace Digital': 93,
'Planet Audio': 94,
'Silicondust': 95,
'Sling Media': 96,
'MEE audio': 97,
'Jensen': 98,
'Goal Zero': 99,
'SANUS': 100,
'Zubie': 101,
'Sandisk': 102,
'Toshiba': 103,
'StarTech': 104,
'ECOXGEAR': 105,
'SOL REPUBLIC': 106,
'House of Marley': 107,
'SiriusXM': 108,
'Panamax': 109,
'Dont': 110

'Logitech': 111,
'Sanus': 112,
'Motorola Home': 113,
'Fitbit': 114,
'Jabra': 115,
'AudioQuest': 116,
'Sengled': 117,
'Cobra Electronics': 118,
'Virgin Mobile': 119,
'VXi': 120,
'Polaroid': 121,
'Linksys': 122,
'Visidec': 123,
'ZAGG': 124,
'Kicker': 125,
'Papago': 126,
'Buffalo Technology': 127,
'Crosley': 128,
'Jaybird': 129,
'Kanto Living': 130,
'Sherwood': 131,
'D-Link': 132,
'House Of Marley': 133,
'Cooler Master': 134,
'Seagate': 135,
'Steelseries': 136,
'Corsair': 137,
'SanDisk': 138,
'Cobra': 139,
'BIC America': 140,
'JBL': 141,
'Tivo': 142,
'ViewSonic': 143,
'FUGOO': 144,
'Pyle': 145,
'Tenba': 146,
'Russound': 147,
'Power Acoustik': 148,
'ONLINE': 149,
'Cerwin Vega': 150,
'CORSAIR': 151,
'Netgear': 152,
'Braven': 153,
'Sima': 154,
'PyleHome': 155,
'Sol Republic': 156,
'Alpine': 157,
'Beats': 158,
'M-Audio': 159,
'MTX Audio': 160,
'StarTech.com': 161,
'Skybell': 162,
'NETGEAR': 163,
'I.am+': 164,
'JVC': 165,
'Arris': 166,
'Jbl': 167,
'V-MODA': 168,
'Cerwin-Vega': 169,
'V-MODA' • 170

'PYLE': 171,
'Kenwood': 172,
'PreSonus': 173,
'VINPOWER DIGITAL DIRECTSHIP': 174,
'Westone': 175,
'KICKER': 176,
'Belkin': 177,
'Escort': 178,
'Klipsch': 179,
'Ultimate Ears': 180,
'Spartan': 181,
'Kanto': 182,
'Western Digital': 183,
'Lenovo': 184,
'Universal Remote Control': 185,
'Pioneer': 186,
'Zoom': 187,
'Bose®': 188,
'DreamWave': 189,
'Rand McNally': 190,
'ZTE': 191,
'Transcend': 192,
'ASUS': 193,
'G-Technology': 194,
'Aiwa': 195,
'NVIDIA': 196,
'Elite Screens': 197,
'Niles': 198,
'SpeakerCraft': 199,
'Garmin': 200,
'Polk Audio': 201,
'RODE': 202,
'Onkyo': 203,
'Mevo': 204,
'Master Dynamic': 205,
'Sonos': 206,
'Definitive Technology': 207,
'TCL': 208,
'Sennheiser': 209,
'Bose': 210,
'Olympus': 211,
'TiVo': 212,
'KEF': 213,
'Buffalo': 214,
'Denon': 215,
'Yamaha': 216,
'360fly': 217,
'WD': 218,
'Nikon': 219,
'Klipsch Xr8i In-Ear Headphones': 220,
'HP': 221,
'Lorex': 222,
'XFX': 223,
'DENON - HEOS': 224,
'Sharp': 225,
'MartinLogan': 226,
'Dell': 227,
'Epson': 228,
'Hisense': 229,
'Mætor & Dynamics': 230

```

'drucker & dynamique': 200,
'Chief': 231,
'Audioquest': 232,
'Panasonic': 233,
'Acer': 234,
'Canon': 235,
'Razer': 236,
'Sigma': 237,
'Evga': 238,
'BenQ': 239,
'CybertronPC': 240,
'Sony': 241,
'VIZIO': 242,
'SVS': 243,
'Apple': 244,
'Samsung': 245,
'Optoma': 246,
'Alienware': 247,
'Asus': 248,
'Marantz': 249,
'MSI': 250,
'Intel': 251,
'LG': 252,
'SunBriteTV': 253}

```

From this list we can see that Sony got rank as 241 w.r.t price
I will create a new column and assign these values

In [62]:

```

dt_electronics['brand_ordinal_labels']=dt_electronics['brand'].map(ordinal_labels2)
dt_electronics.head()

```

Out[62]:

	pricesavailability	pricescondition	pricesdateSeen	pricesmerchant	pricesshipping
0	1590	4645	2017-05-10T20:00:00Z,2017-05-09T15:00:00Z	2109	
1	1590	4645	2017-10-10T02:00:00Z,2017-08-12T03:00:00Z,2017...	2109	
2	1590	4645	2017-10-10T19:00:00Z,2017-09-12T14:00:00Z,2017...	2109	
3	1590	4645	2017-09-08T05:00:00Z,2017-09-18T13:00:00Z,2017...	2109	

```
4          63      4645      2017-12-  
05T13:00:00Z      1119
```

We have two date field dateUpdated,dateAdded.First I will convert them to datetime

In [63]:

```
dt_electronics['dateUpdated'] = pd.to_datetime(dt_electronics['dateUpdated']).astype('datetime64[ns]',format='%y%m%d')
```

In [64]:

```
dt_electronics['dateAdded'] = pd.to_datetime(dt_electronics['dateAdded']).astype('datetime64[ns]',format='%y%m%d')
```

Now i create month and year column from this date field

In [65]:

```
dt_electronics['Month_dtadd']=pd.DatetimeIndex(dt_electronics['dateAdded']).month  
dt_electronics['Year_dtadd']=pd.DatetimeIndex(dt_electronics['dateAdded']).year
```

In [66]:

```
dt_electronics['Month_updated']=pd.DatetimeIndex(dt_electronics['dateUpdated']).month  
dt_electronics['Year_updated']=pd.DatetimeIndex(dt_electronics['dateUpdated']).year
```

In [67]:

```
dt_electronics.head()
```

Out[67]:

	pricesavailability	pricescondition	pricesdateSeen	pricesmerchant	pricesshipping
0	1590	4645	2017-05-05T20:00:00Z,2017-05-09T15:00:00Z	2109	
1	1590	4645	2017-10-08T02:00:00Z,2017-12T03:00:00Z,2017...	2109	

2	1590	4645	2017-10- 10T19:00:00Z,2017- 09- 12T14:00:00Z,2017...	2109
3	1590	4645	2017-09- 08T05:00:00Z,2017- 09- 18T13:00:00Z,2017...	2109
4	63	4645	2017-12- 05T13:00:00Z	1119

Now I will drop this two columns because I have extracted all the relevant information from these fields

In [68]:

```
dt_electronics.drop(['dateAdded', 'dateUpdated'], axis=1, inplace=True)
```

In [69]:

```
dt_electronics.head()
```

Out[69]:

	pricesavailability	pricescondition	pricesdateSeen	pricesmerchant	pricesshipping
0	1590	4645	2017-05- 10T20:00:00Z,2017- 05-09T15:00:00Z	2109	
1	1590	4645	2017-10- 10T02:00:00Z,2017- 08- 12T03:00:00Z,2017...	2109	
2	1590	4645	2017-10- 10T19:00:00Z,2017- 09- 12T14:00:00Z,2017...	2109	
3	1590	4645	2017-09- 08T05:00:00Z,2017- 09- 18T13:00:00Z,2017...	2109	
4	63	4645	2017-12- 05T13:00:00Z	1119	

```
In [70]:
```

```
dt_electronics.manufacturer.unique()
```

```
Out[70]:
```

```
array(['Ecoxgear', nan, 'House Of Marley', 'Sony', 'Acer', 'Yamaha',
       'KICKER', 'Russound', 'Manufacturers Limited Warranty: Lifetime',
       'Samsung', 'Manufacturers Limited Warranty:3 Year',
       'House of Marley', 'Joby', 'Sonic Alert', '120', 'Elite Screens',
       'RODE', 'Audiovox', 'Alpine', 'Razer', 'Polk Audio',
       'Pioneer',
       'StarTech.com', 'JVC', 'Canon Cameras US', 'Samsung IT',
       'Ilive Blue', 'Seagate Technology', 'Apple', 'Power Acoustik',
       'Yamaha Electronics', 'Polaroid', 'iSimple', 'Dell',
       'Westone Laboratories, Inc.', 'Kanto', 'Panasonic', 'Canon',
       'Corsair', 'Peerless', 'Zoom', 'Sennheiser', 'DenAQ',
       'Bose',
       'Western Digital', 'Slingbox', 'Sima', 'Pac', 'NETGEAR',
       'Manfrotto', 'Zmodo', 'Vizio',
       'Cooks International LLC (Former Coshell Holdings LLC)', 'JBL',
       'Boytone', 'SiriusXM', 'Grace Digital Audio', 'Apple Computer',
       'Microsoft', 'Dual', 'ZTE', 'V-Moda',
       'Manufacturers Limited Warranty:1 Year', 'Cerwin Vega',
       'NVIDIA',
       'Allround Software', 'Sanus', 'Samsung Electronics',
       'Definitive Technology', 'Parts', 'Logitech', 'Asus',
       'Lowepro',
       'G-TECH', 'Denon', 'Actiontec', 'SiliconDust USA, Inc',
       'TP-Link',
       'Midland', 'Bower', 'Sabrent',
       'Klipsch X12 Bluetooth Neckband Headphones (black Leather)',
       'ONKYO', 'Manufacturers Limited Warranty:2 Year', 'Sharp',
       'KENWOOD', 'SanDisk', 'BIC America', 'Incipio', 'Logitech, Inc',
       'SAMSUNG', 'PYLE', 'Limited Lifetime Warranty', 'Sylvania',
       'Garmin', 'Jensen', 'Belkin', 'Lenovo Idea',
       'Viper by GLD Products', 'AudioQuest', 'Case Logic',
       '1 Year',
       'BOSE', 'iHome', 'Buffalo Technology', 'ISIMPLE', 'Elite Screens',
       'GoPro', 'Kenwood',
       'Jbl Everest 110 In-Ear Wireless Bluetooth Headphones (silver)',
       'Lenovo Corporate', 'TiVo', 'RCA', 'Grace Digital', 'Beats',
       'Sound Around', 'Tiffen', 'KT.TPSCH', '2 Years'])
```

```
'Apple Computer (Direct)', 'Netgear Inc', 'Coolermaster',
 'Sherwood', 'Peerless Industries', 'Epson', 'Sigma',
 'Zagg',
 'Mophie', 'MTX', 'Speck', '1 year warranty',
 'Sony Mobile Communications, (USA) Inc', 'Novatel', 'LG ODD',
 '5 Years', 'OtterBox', 'YAMAHA', 'Bower Camera', 'Onkyo',
 'ARRIS Solutions, Inc.', 'Aluratek', 'Pelican',
 'EPSON AMERICA, INC.', 'Visiontek', 'Audioquest', 'Spartan',
 'RAZER', 'Boyo', 'TRANSCEND Storage', 'JAYBIRD',
 'Urban Armor Gear', 'NZXT', 'SKB Cases', 'Planet Audio',
 'Fugoo',
 'PreSonus', 'Peak Design', 'Fusion', 'Go-Pro', 'Marantz',
 'Sunpak',
 'Spacepole Inc', 'Siriusxm', 'Atrend', 'HP Inc', 'MEE audio',
 'Papago', 'Sonos', 'ELITE SCREENS DIRECTSHIP', 'AUDIOQUEST',
 'x6d inc', 'ECO STYLE', 'StarTech', 'Kicker', '1-Year',
 'Cybertronpc', 'Chief Manufacturing', 'Diamond Multimedia',
 'Pyle',
 'Alpine Electronics USA', 'Razer USA', 'niles', 'Lorex',
 'Crosley Radio', 'Thule', 'MTRLC LLC',
 'Jvc Real Sound System Z Series Ha-Sz2000 Stream Woofers Db Headphones (black, Japan Import) V-Moda',
 'Klipsch', 'Sonax', 'eTape16', 'Motorola', 'Dantona',
 'OmniMount',
 'LowePro', 'Niles', 'Lenovo', 'Energizer', 'Emerge'],
 dtype=object)
```

In [71]:

```
dt_electronics.manufacturer.value_counts()
```

Out[71]:

```
Sony
179
Samsung
132
Apple
101
Pioneer
78
Yamaha
52
Elitescreens
46
Seagate Technology
41
KENWOOD
38
Alpine
37
Corsair
--
```

35
Lowepro
33
Onkyo
33
Russound
31
V-Moda
27
Epson
27
Sennheiser
26
Denon
25
Actiontec
24
Western Digital
24
Garmin
24
Parts
24
Logitech
23
Canon
23
YAMAHA
23
Samsung Electronics
22
NETGEAR
21
Zoom
21
Dell
21
ONKYO
20
Denaq
20
Panasonic
20
Acer
19
120
18
SiriusXM
18
Power Acoustik
18
Manufacturers Limited Warranty:2 Year
18
House of Marley
18
Yamaha Electronics
17
Joby
16
Grace Digital
15

✓ Samsung IT
15
Kanto
15
iSimple
15
Manufacturers Limited Warranty:3 Year
15
Asus
14
Sylvania
14
MTX
13
Manufacturers Limited Warranty:Lifetime
13
Grace Digital Audio
13
StarTech.com
13
Razer
13
SiliconDust USA, Inc
13
Bose
13
Sanus
13
Manfrotto
13
Papago
13
SanDisk
13
Kenwood
12
JVC
12
Vizio
12
Logitech, Inc
12
Belkin
12
Lenovo Corporate
12
Boytone
11
RAZER
11
iHome
11
OtterBox
11
TP-Link
11
NZXT
11
Jensen
11

++
BIC America
11
Manufacturers Limited Warranty: 1 Year
11
Ecoxgear
11
Audioquest
11
Westone Laboratories, Inc.
11
LG ODD
10
RCA
10
Sony Mobile Communications, (USA) Inc
10
KICKER
10
House Of Marley
10
ISIMPLE
10
SAMSUNG
10
Sima
10
Sherwood
9
Netgear Inc
9
RODE
9
Canon Cameras US
9
Audiovox
9
Jbl Everest 110 In-Ear Wireless Bluetooth Headphones (silver)
8
Midland
8
Sharp
8
Zagg
8
Dual
8
AudioQuest
8
JAYBIRD
8
1 year warranty
8
Tiffen
8
Cybertronpc
7
MEE audio
7
Viper by GLD Products
7

'
Sigma
7
PYLE
7
Peak Design
7
Visiontek
7
JBL
7
Lenovo Idea
7
Apple Computer
7
Spacepole Ince
7
Peerless
7
Pyle
7
Microsoft
7
Aluratek
6
Pelican
6
Zmodo
6
SKB Cases
6
Polaroid
6
Klipsch
6
LowePro
6
Spartan
6
MTRLC LLC
6
Apple Computer (Direct)
6
Speck
6
Limited Lifetime Warranty
6
ELITE SCREENS DIRECTSHIP
6
BOSE
6
5 Years
6
KLIPSCH
6
1 Year
6
Coolermaster
6
Allround Software
6

^
Elite Screens
6
ARRIS Solutions, Inc.
6
Incipio
5
Buffalo Technology
5
Marantz
5
Sound Around
5
eTape16
5
Bower Camera
5
TiVo
5
Urban Armor Gear
5
Boyo
5
Ilive Blue
5
Novatel
5
Sonos
5
StarTech
4
2 Years
4
Thule
4
EPSON AMERICA, INC.
4
Planet Audio
4
Siriusxm
4
TRANSCEND Storage
4
Pac
4
Sonic Alert
4
Fugoo
4
Fusion
4
Bower
4
Emerge
4
Klipsch X12 Bluetooth Neckband Headphones (black Leather)
4
Dantona
4
Jvc Real Sound System Z Series Ha-Sz2000 Stream Woofer Db Hea
dphones (black Tannen Transport) V-Modes

APPLIED MATERIALS, INC., V. SONY
Cerwin Vega
4
Go-Pro
4
Mophie
4
Peerless Industries
3
Crosley Radio
3
ZTE
3
GoPro
3
OmniMount
3
x6d inc
3
ECO STYLE
3
niles
3
Atrend
3
HP Inc
3
Sabrent
3
G-TECH
3
Beats
3
Slingbox
3
PreSonus
3
Razer USA
2
Cooks International LLC (Former Coshell Holdings LLC)
2
Polk Audio
2
1-Year
2
Alpine Electronics USA
2
Sunpak
2
Niles
2
NVIDIA
2
Motorola
2
Kicker
2
Chief Manufacturing
2
Definitive Technology
2

```
Case Logic  
2  
Diamond Multimedia  
1  
Lorex  
1  
Sonax  
1  
Lenovo  
1  
AUDIOQUEST  
1  
Energizer  
1  
Name: manufacturer, dtype: int64
```

I will calculate the count of each unique values of manufacturer and create a new column

In [72]:

```
manufacturer_map=dt_electronics['manufacturer'].value_counts().to_dict()  
manufacturer_map
```

Out[72]:

```
{'Sony': 179,  
'Samsung': 132,  
'Apple': 101,  
'Pioneer': 78,  
'Yamaha': 52,  
'Elitescreens': 46,  
'Seagate Technology': 41,  
'KENWOOD': 38,  
'Alpine': 37,  
'Corsair': 35,  
'Lowepro': 33,  
'Onkyo': 33,  
'Russound': 31,  
'V-Moda': 27,  
'Epson': 27,  
'Sennheiser': 26,  
'Denon': 25,  
'Actiontec': 24,  
'Western Digital': 24,  
'Garmin': 24,  
'Parts': 24,  
'Logitech': 23,  
'Canon': 23,  
'YAMAHA': 23,  
'Samsung Electronics': 22,  
'NETGEAR': 21,  
'Zoom': 21,  
'Dell': 21,  
'ONKYO': 20,  
'Denaa': 20.}
```

'Panasonic': 20,
'Acer': 19,
'120': 18,
'SiriusXM': 18,

'Power Acoustik': 18,
'Manufacturers Limited Warranty:2 Year': 18,
'House of Marley': 18,
'Yamaha Electronics': 17,
'Joby': 16,
'Grace Digital': 15,
'Samsung IT': 15,
'Kanto': 15,
'iSimple': 15,
'Manufacturers Limited Warranty:3 Year': 15,
'Asus': 14,
'Sylvania': 14,
'MTX': 13,
'Manufacturers Limited Warranty:Lifetime': 13,
'Grace Digital Audio': 13,
'StarTech.com': 13,
'Razer': 13,
'SiliconDust USA, Inc': 13,
'Bose': 13,
'Sanus': 13,
'Manfrotto': 13,
'Papago': 13,
'SanDisk': 13,
'Kenwood': 12,
'JVC': 12,
'Vizio': 12,
'Logitech, Inc': 12,
'Belkin': 12,
'Lenovo Corporate': 12,
'Boytone': 11,
'RAZER': 11,
'iHome': 11,
'OtterBox': 11,
'TP-Link': 11,
'NZXT': 11,
'Jensen': 11,
'BIC America': 11,
'Manufacturers Limited Warranty:1 Year': 11,
'Ecoxgear': 11,
'Audioquest': 11,
'Westone Laboratories, Inc.': 11,
'LG ODD': 10,
'RCA': 10,
'Sony Mobile Communications, (USA) Inc': 10,
'KICKER': 10,
'House Of Marley': 10,
'ISIMPLE': 10,
'SAMSUNG': 10,
'Sima': 10,
'Sherwood': 9,
'Netgear Inc': 9,
'RODE': 9,
'Canon Cameras US': 9,
'Audiovox': 9,
'Jbl Everest 110 In-Ear Wireless Bluetooth Headphones (silver)': 9

'Midland': 8,
'Sharp': 8,
'Zagg': 8,
'Dual': 8,

'AudioQuest': 8,
'JAYBIRD': 8,
'1 year warranty': 8,
'Tiffen': 8,
'Cybertronpc': 7,
'MEE audio': 7,
'Viper by GLD Products': 7,
'Sigma': 7,
'PYLE': 7,
'Peak Design': 7,
'Visiontek': 7,
'JBL': 7,
'Lenovo Idea': 7,
'Apple Computer': 7,
'Spacepole Ince': 7,
'Peerless': 7,
'Pyle': 7,
'Microsoft': 7,
'Aluratek': 6,
'Pelican': 6,
'Zmodo': 6,
'SKB Cases': 6,
'Polaroid': 6,
'Klipsch': 6,
'LowePro': 6,
'Spartan': 6,
'MTRLC LLC': 6,
'Apple Computer (Direct)': 6,
'Speck': 6,
'Limited Lifetime Warranty': 6,
'ELITE SCREENS DIRECTSHIP': 6,
'BOSE': 6,
'5 Years': 6,
'KLIPSCH': 6,
'1 Year': 6,
'Coolermaster': 6,
'Allround Software': 6,
'Elite Screens': 6,
'ARRIS Solutions, Inc.': 6,
'Incipio': 5,
'Buffalo Technology': 5,
'Marantz': 5,
'Sound Around': 5,
'eTape16': 5,
'Bower Camera': 5,
'TiVo': 5,
'Urban Armor Gear': 5,
'Boyo': 5,
'Ilive Blue': 5,
'Novatel': 5,
'Sonos': 5,
'StarTech': 4,
'2 Years': 4,
'Thule': 4,
'EPSON AMERICA, INC.': 4,
'Planet Audio': 4

```
        'Planar Audio': 2,
        'Siriusxm': 4,
        'TRANSCEND Storage': 4,
        'Pac': 4,
        'Sonic Alert': 4,
        'Fugoo': 4,
        'Fusion': 4,
        'Bower': 4,
        'Emerge': 4,
        'Klipsch X12 Bluetooth Neckband Headphones (black Leather)': 4,
        'Dantona': 4,
        'Jvc Real Sound System Z Series Ha-Sz2000 Stream Woofer Db Headphones (black, Japan Import) V-Moda': 4,
        'Cerwin Vega': 4,
        'Go-Pro': 4,
        'Mophie': 4,
        'Peerless Industries': 3,
        'Crosley Radio': 3,
        'ZTE': 3,
        'GoPro': 3,
        'OmniMount': 3,
        'x6d inc': 3,
        'ECO STYLE': 3,
        'niles': 3,
        'Atrend': 3,
        'HP Inc': 3,
        'Sabrent': 3,
        'G-TECH': 3,
        'Beats': 3,
        'Slingbox': 3,
        'PreSonus': 3,
        'Razer USA': 2,
        'Cooks International LLC (Former Coshell Holdings LLC)': 2,
        'Polk Audio': 2,
        '1-Year': 2,
        'Alpine Electronics USA': 2,
        'Sunpak': 2,
        'Niles': 2,
        'NVIDIA': 2,
        'Motorola': 2,
        'Kicker': 2,
        'Chief Manufacturing': 2,
        'Definitive Technology': 2,
        'Case Logic': 2,
        'Diamond Multimedia': 1,
        'Lorex': 1,
        'Sonax': 1,
        'Lenovo': 1,
        'AUDIOQUEST': 1,
        'Energizer': 1}
```

here I will replace 'Apple' with 101

In [73]:

```
dt_electronics['manufacturer_count']=dt_electronics['manufacturer'].map(ma
```

```
nufacturer_map)
dt_electronics.head()
```

Out[73]:

	pricesavailability	pricescondition	pricesdateSeen	pricesmerchant	pricesshipping
0	1590	4645	2017-05-05T20:00:00Z,2017-05-09T15:00:00Z	2109	
1	1590	4645	2017-10-08T02:00:00Z,2017-12T03:00:00Z,2017...	2109	
2	1590	4645	2017-10-09T19:00:00Z,2017-12T14:00:00Z,2017...	2109	
3	1590	4645	2017-09-08T05:00:00Z,2017-18T13:00:00Z,2017...	2109	
4	63	4645	2017-12-05T13:00:00Z	1119	

In [74]:

```
dt_electronics.columns
```

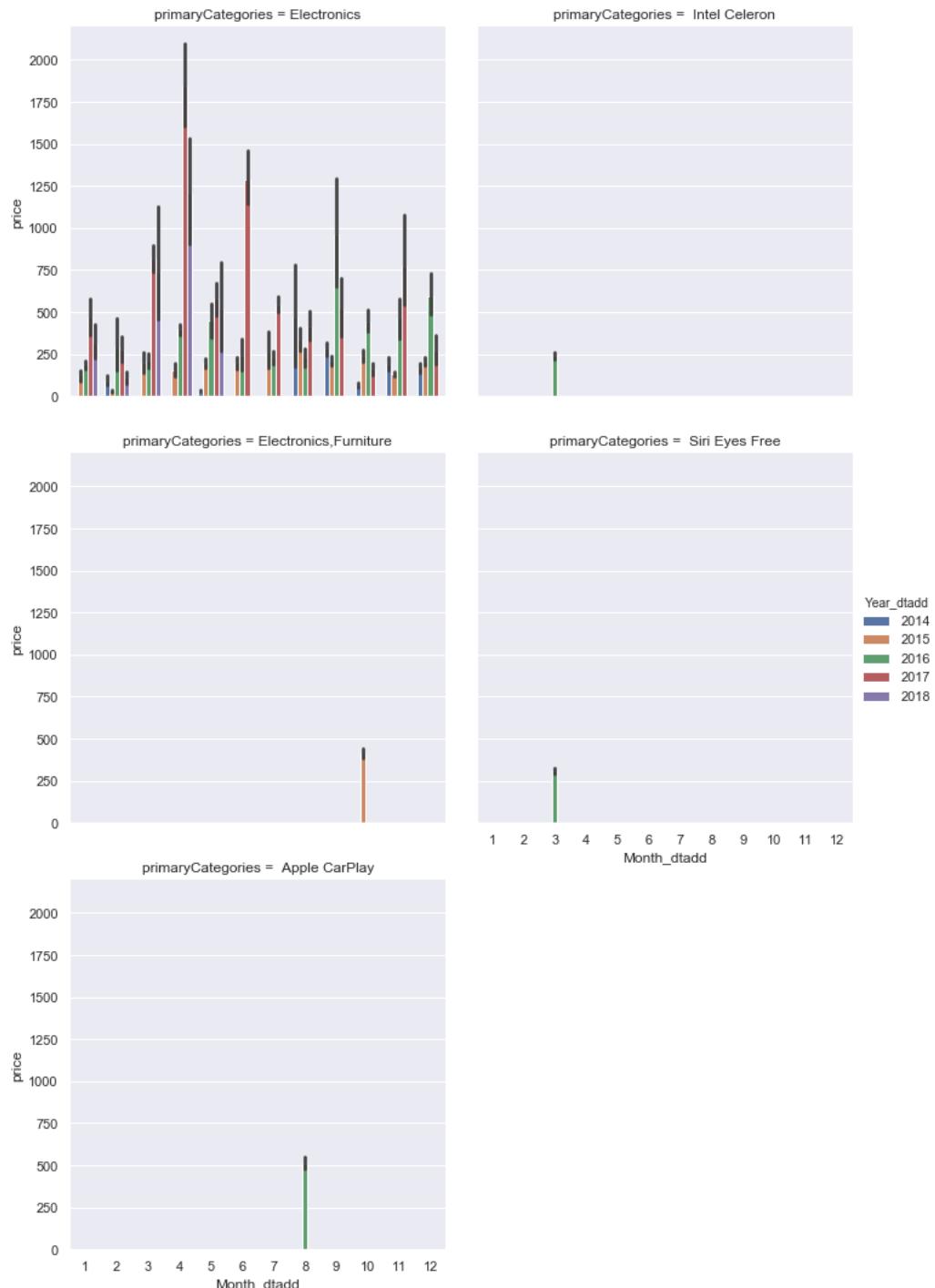
Out[74]:

```
Index(['pricesavailability', 'pricescondition', 'pricesdateSeen',
       'pricesmerchant', 'pricesshipping', 'pricessourceURLs', 'asins',
       'brand', 'categories', 'ean', 'imageURLs', 'keys', 'manufacturer',
       'manufacturerNumber', 'name', 'primaryCategories', 'sourceURLs',
       'upc',
       'weight', 'price', 'Currency_USD', 'pricesisSale_True',
       'brand_ordinal_labels', 'Month_dtadd', 'Year_dtadd',
       'Month_updated',
       'Year_updated', 'manufacturer_count'],
      dtype='object')
```

I will plot primaryCategories with dateAdded

In [71]:

```
sns.set(style='darkgrid', font_scale=1.0)
sns.catplot(x='Month_dtadd',
            y='price', hue='Year_dtadd',
            col='primaryCategories',
            col_wrap=2,
            data=dt_electronics,
            kind='bar')
plt.savefig('Dateadded_primaryCategories.png')
```



Observation

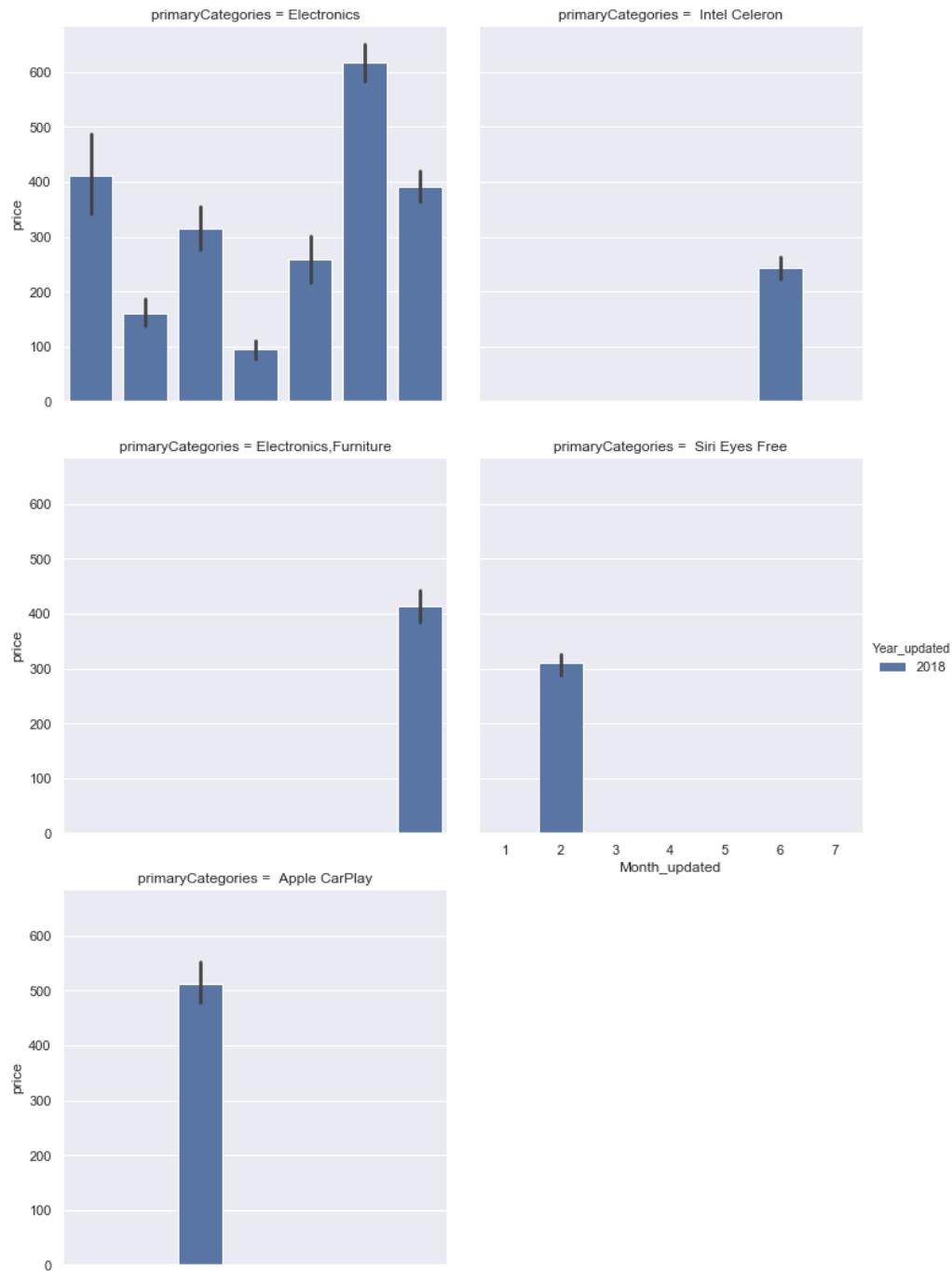
From this graph we can conclude Electronics category has more price than other category

I can also conclude that in 4th month of 2017 price is highest

I will plot primaryCategories with dateUpdated

In [72]:

```
sns.set(style='darkgrid', font_scale=1.0)
sns.catplot(x='Month_updated',
            y='price',hue='Year_updated',
            col='primaryCategories',
            col_wrap=2,
            data=dt_electronics,
            kind='bar')
plt.savefig('Dateupdated_primaryCategories.png')
```





Observation

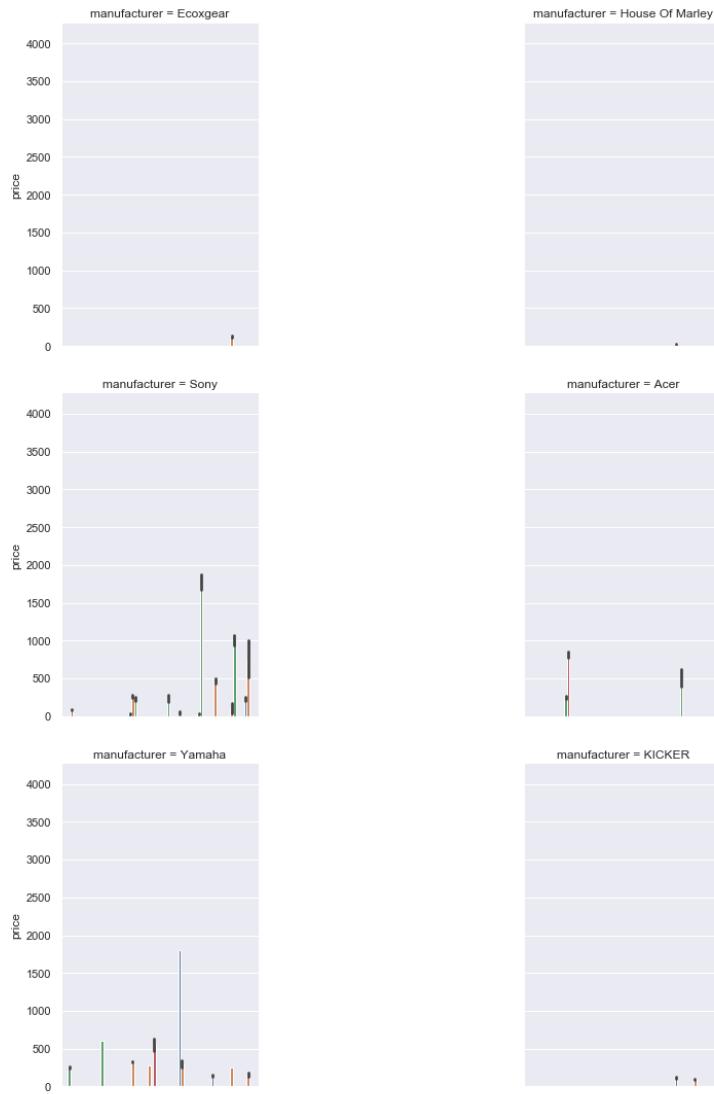
From this graph we can conclude Electronics category has more price than other category

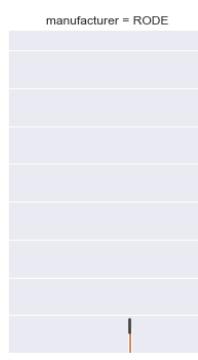
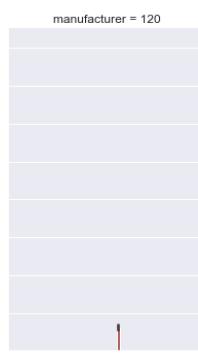
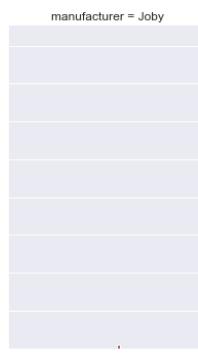
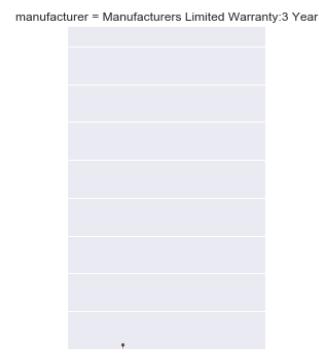
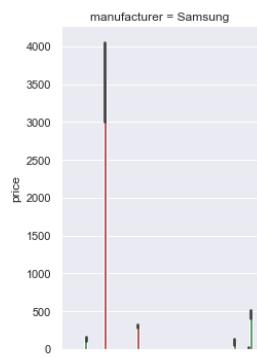
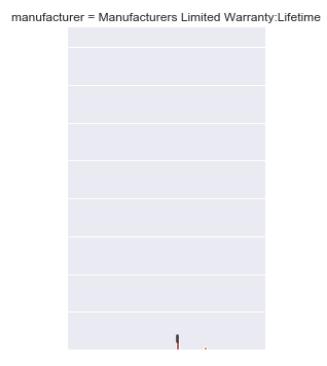
I can also conclude that in 6th month of 2018 price is highest

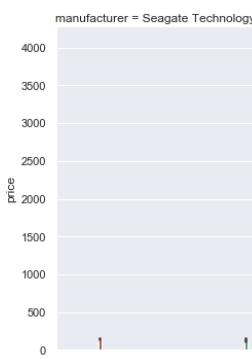
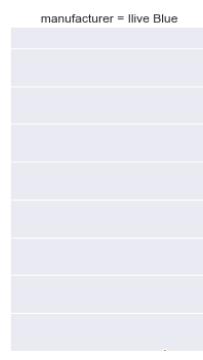
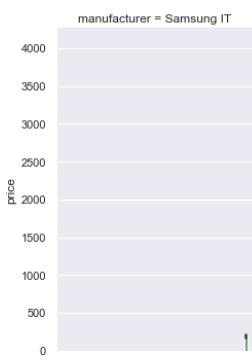
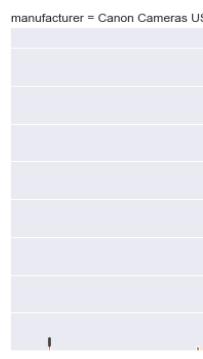
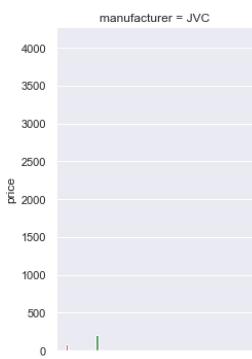
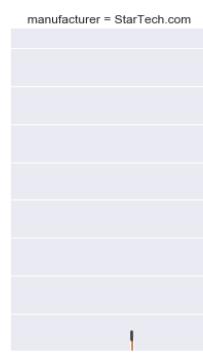
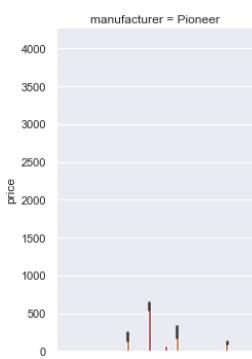
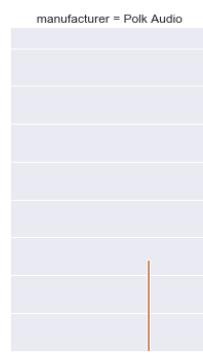
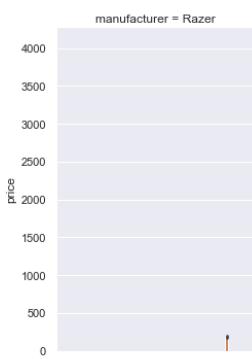
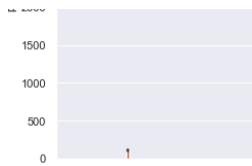
I will plot manufacturer with dateAdded

In [73]:

```
sns.set(style='darkgrid', font_scale=1.0)
sns.catplot(x='Month_dtadd',
            y='price', hue='Year_dtadd',
            col='manufacturer',
            col_wrap=2,
            data=dt_electronics,
            kind='bar')
plt.savefig('Dateadded_manufacture.png')
```

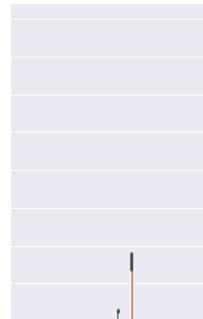




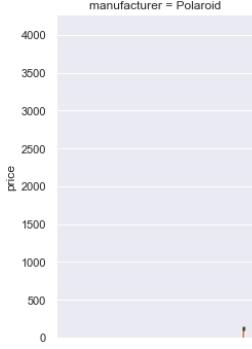




manufacturer = Polaroid



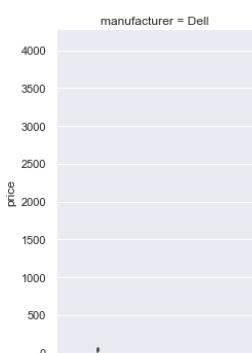
manufacturer = iSimple



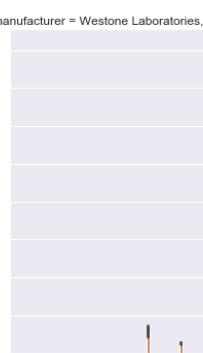
manufacturer = Dell



manufacturer = Westone Laboratories, Inc.



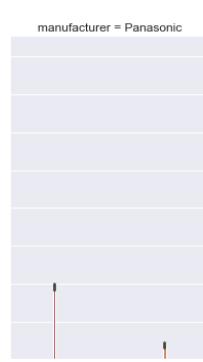
manufacturer = Kanto



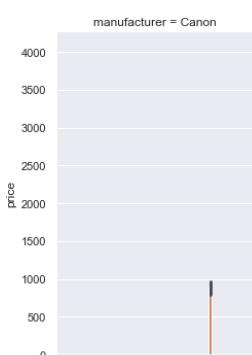
manufacturer = Panasonic



manufacturer = Canon



manufacturer = Corsair



manufacturer = Peerless



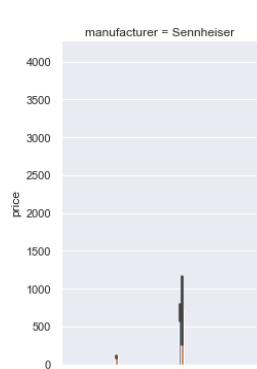
manufacturer = Zoom



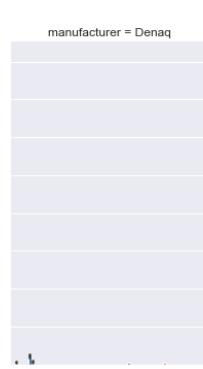
manufacturer = Sennheiser



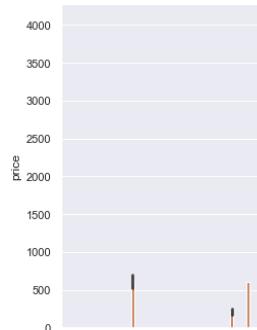
manufacturer = Denaq



manufacturer = Bose



manufacturer = Western Digital



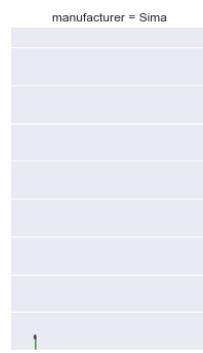
manufacturer = Slingbox



manufacturer = Sima



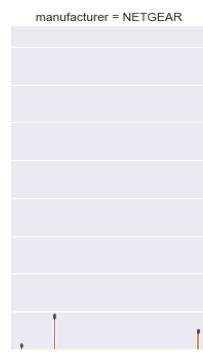
manufacturer = Pac



manufacturer = NETGEAR



manufacturer = Manfrotto



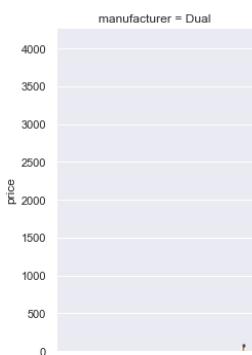
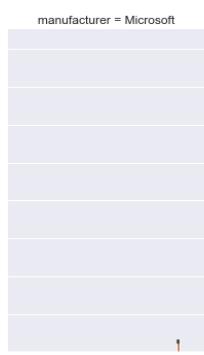
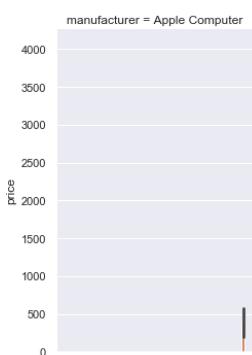
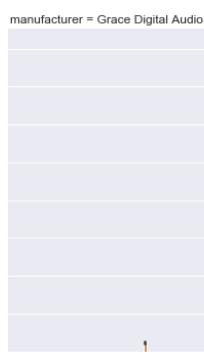
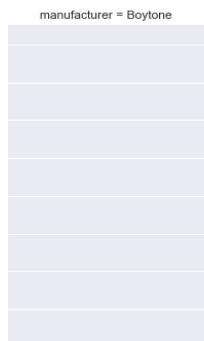
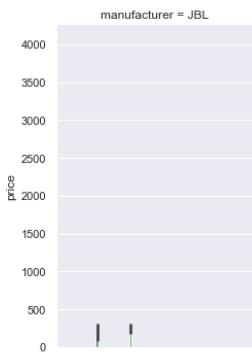
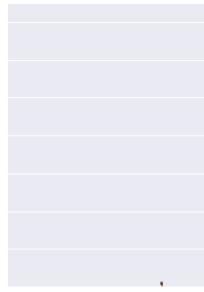
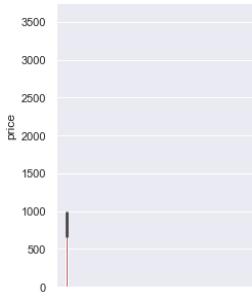
manufacturer = Zmodo

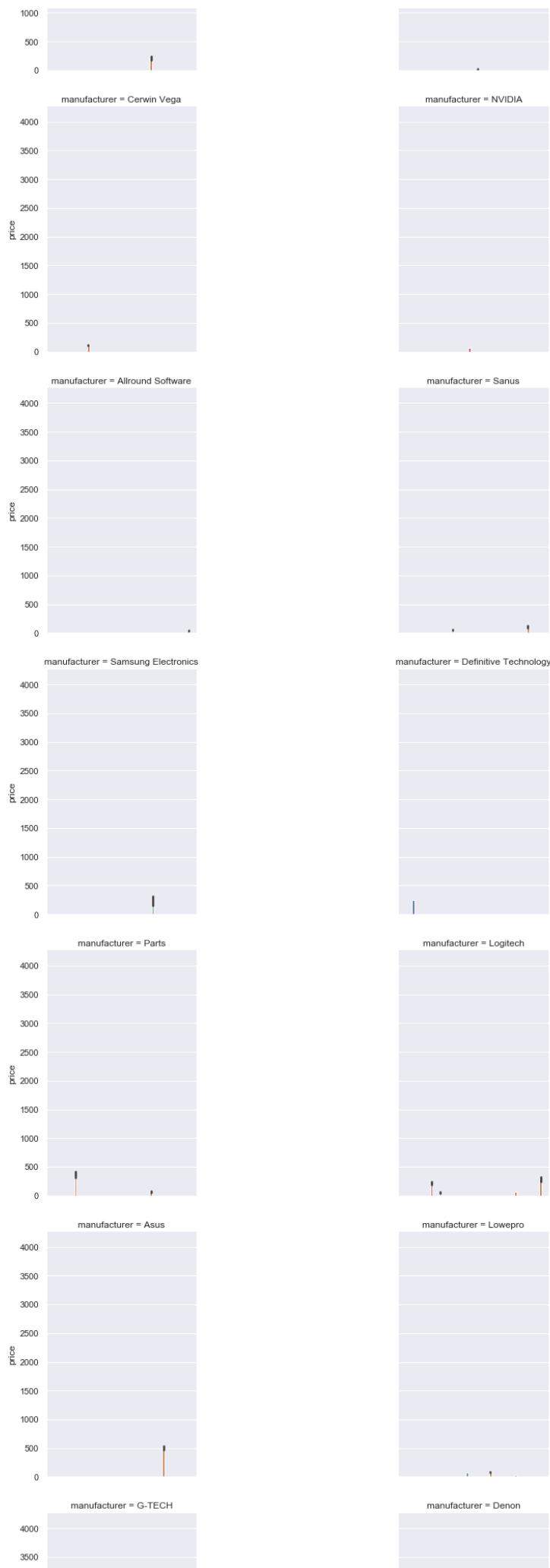


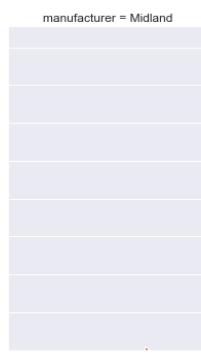
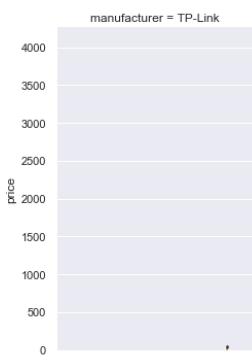
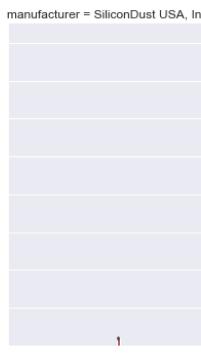
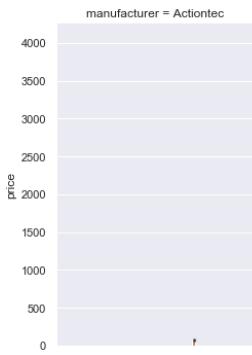
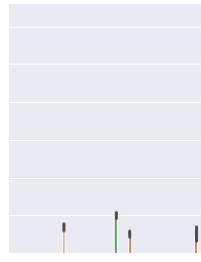
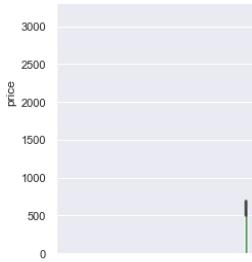
manufacturer = Vizio

manufacturer = Cooks International LLC (Former Coshell Holdings LLC)

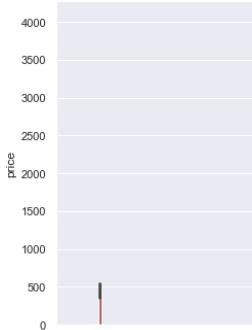




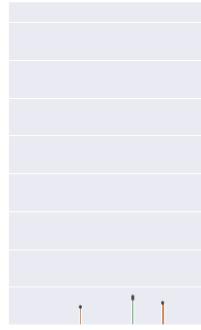




manufacturer = Klipsch X12 Bluetooth Neckband Headphones (black Leather)



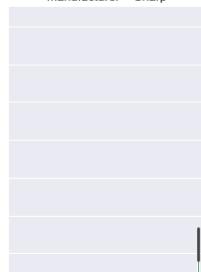
manufacturer = ONKYO

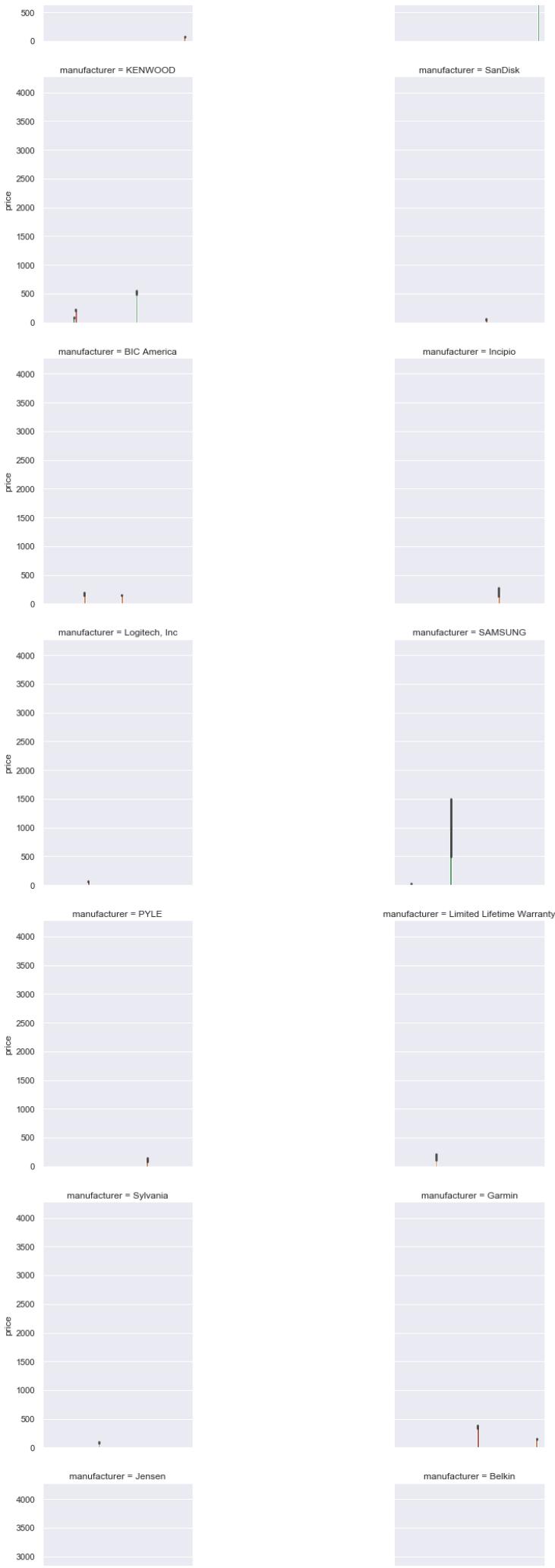


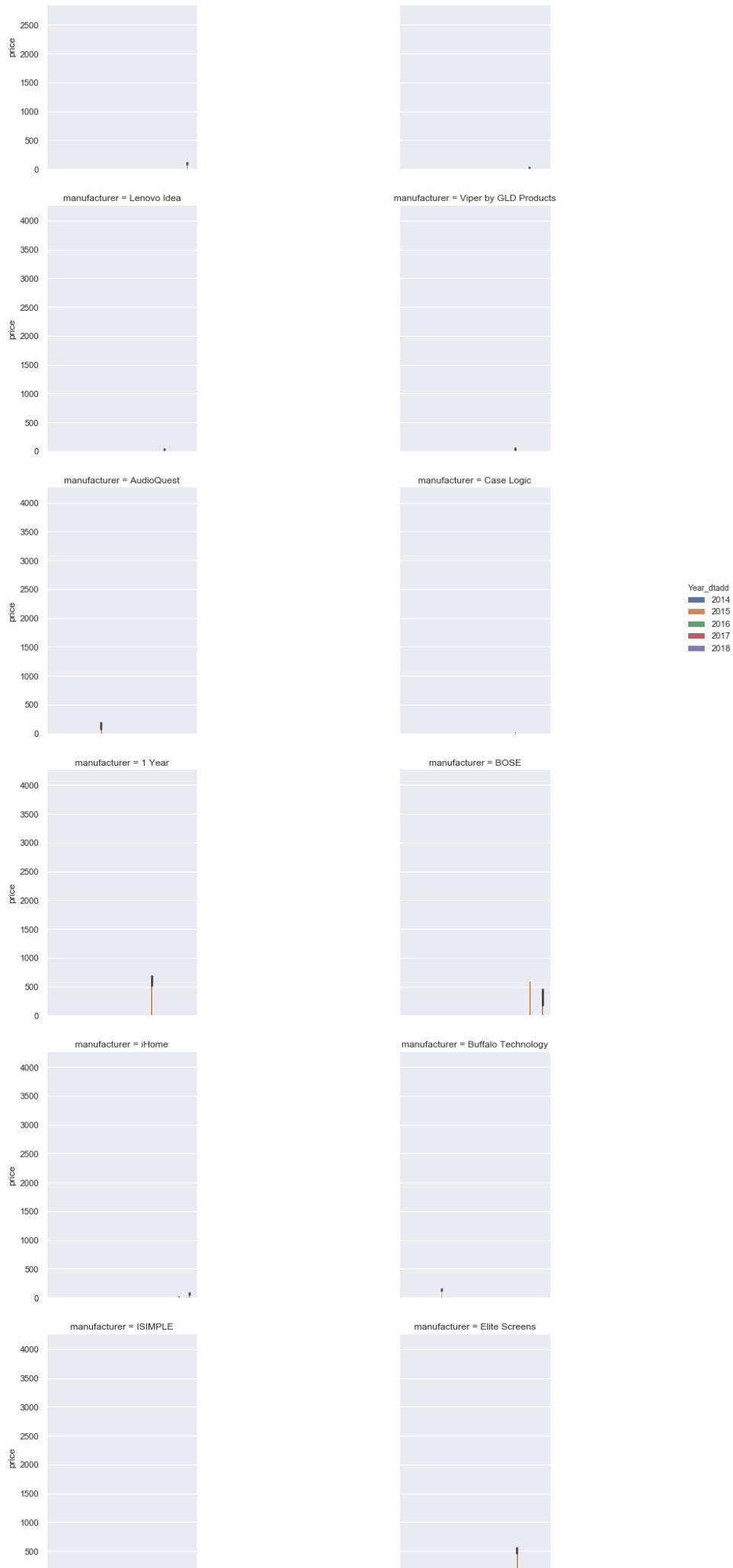
manufacturer = Manufacturers Limited Warranty:2 Year



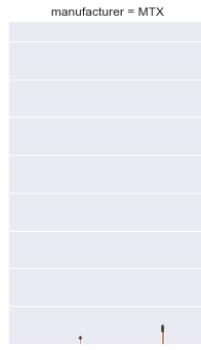
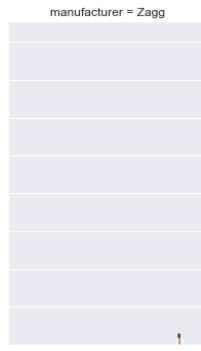
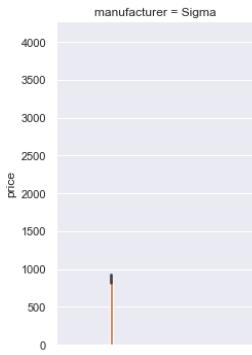
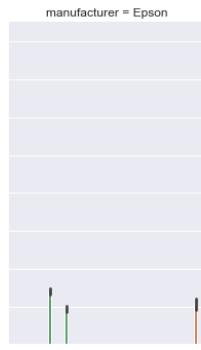
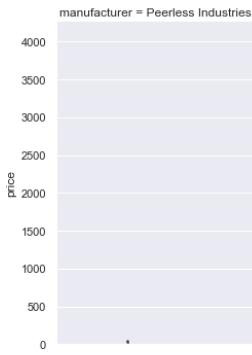
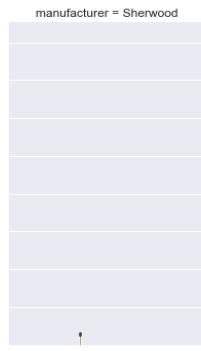
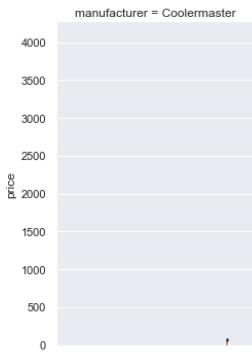
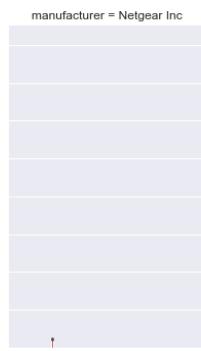
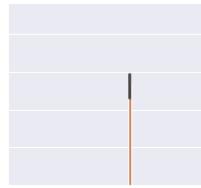
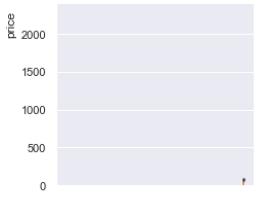
manufacturer = Sharp

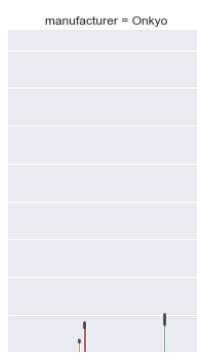
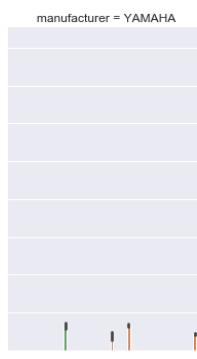
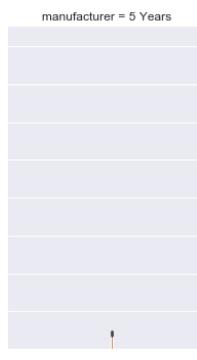
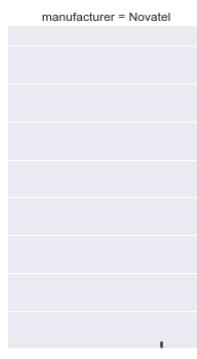
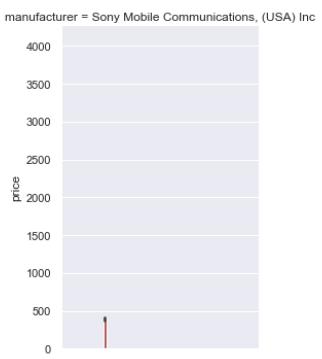
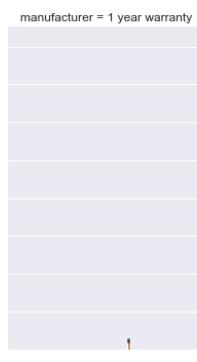














manufacturer = Pelican



manufacturer = EPSON AMERICA, INC.

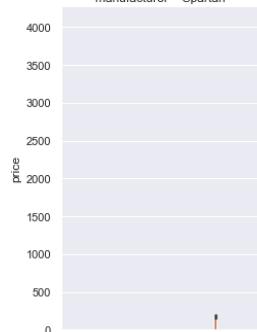


manufacturer = Visiontek

manufacturer = Audioquest



manufacturer = Audioquest



manufacturer = Spartan

manufacturer = RAZER



manufacturer = Boyo

manufacturer = TRANSCEND Storage



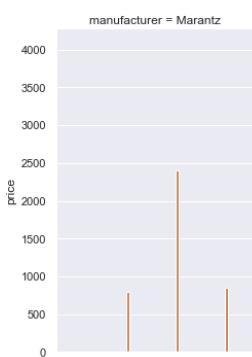
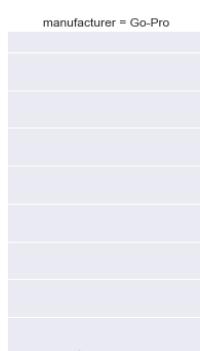
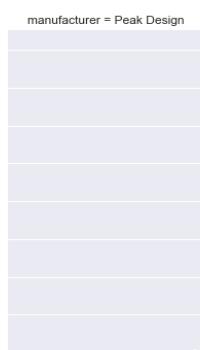
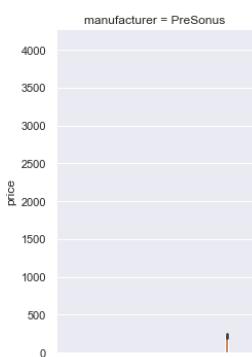
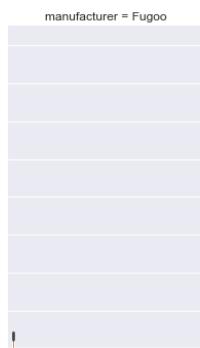
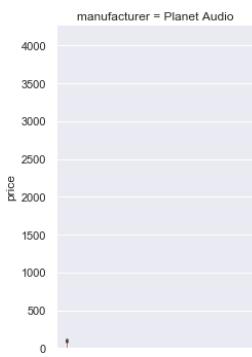
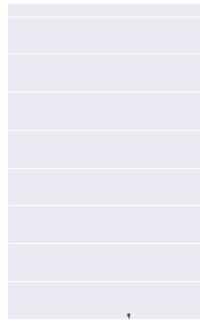
manufacturer = JAYBIRD

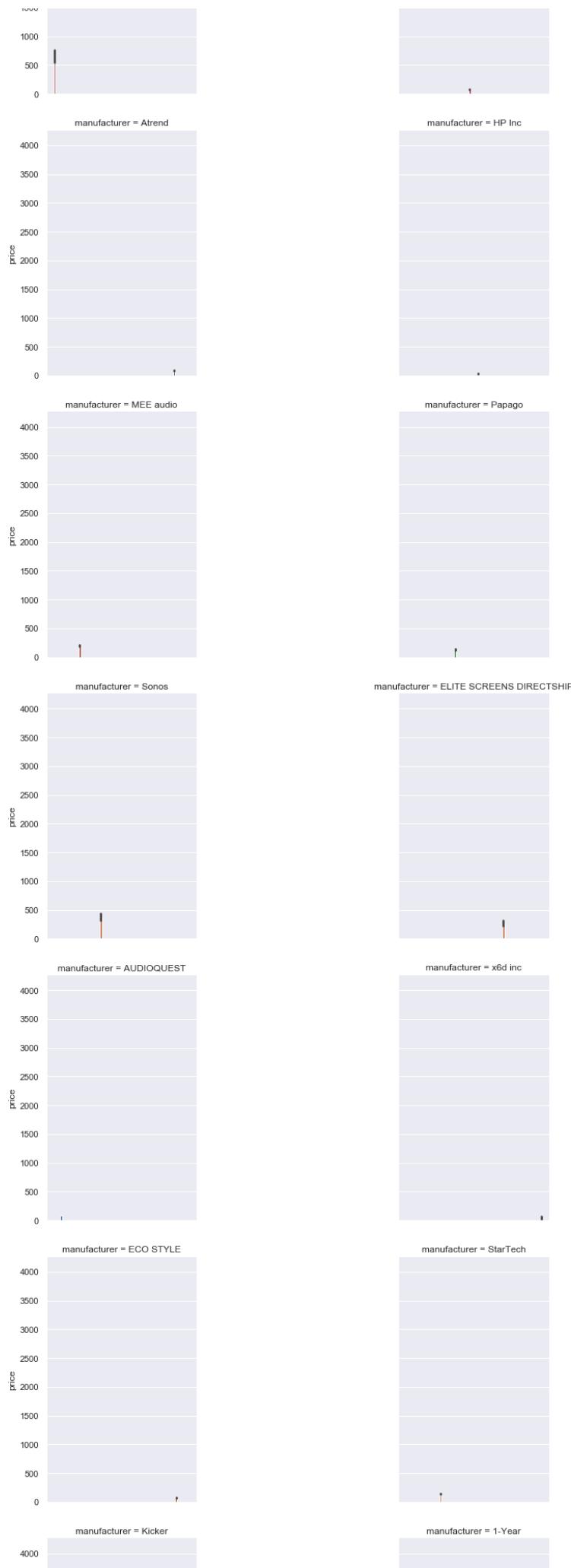
manufacturer = Urban Armor Gear

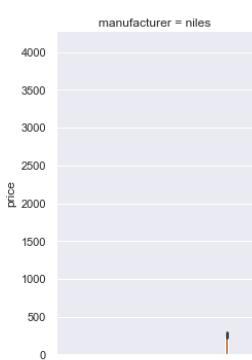
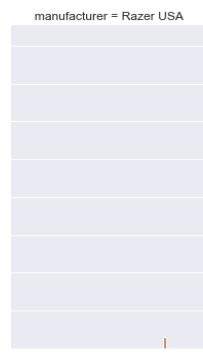
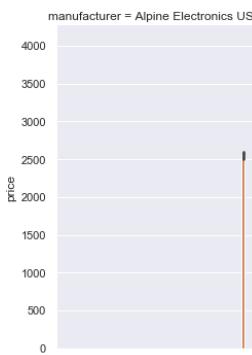
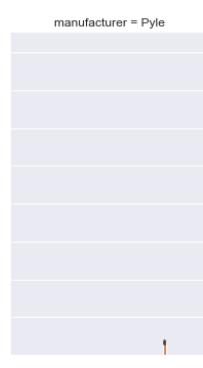
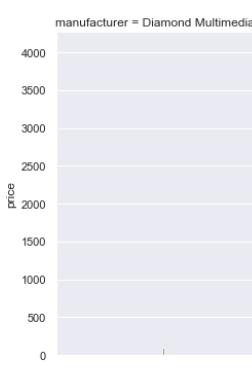
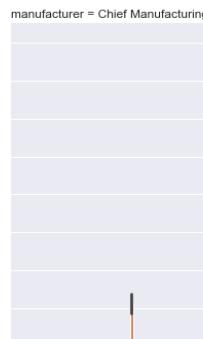
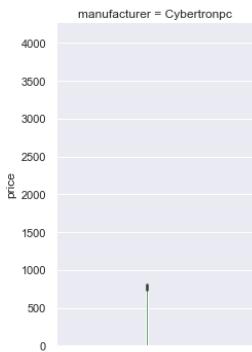
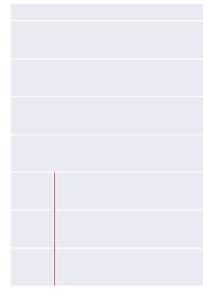


manufacturer = NZXT

manufacturer = SKB Cases

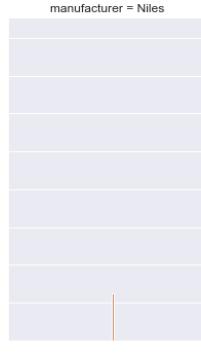
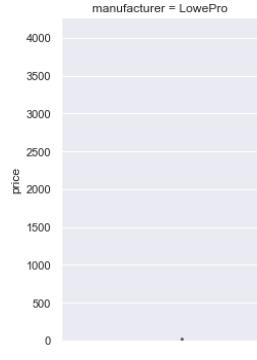
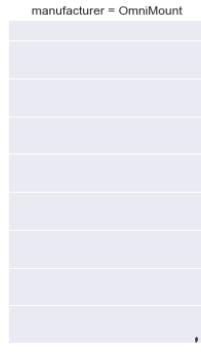
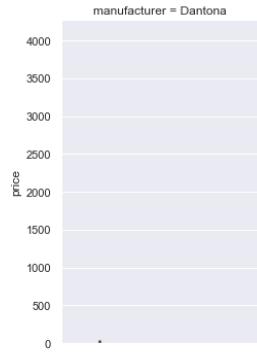
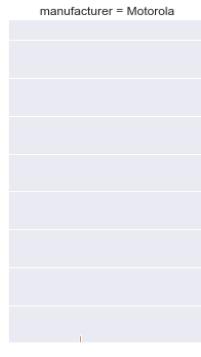
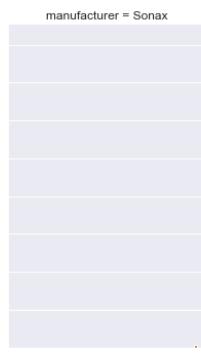
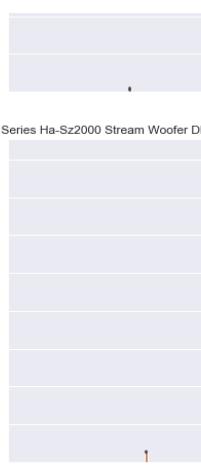








manufacturer = Jvc Real Sound System Z Series Ha-Sz2000 Stream Woofer Db Headphones (black, Japan Import) V-Moda





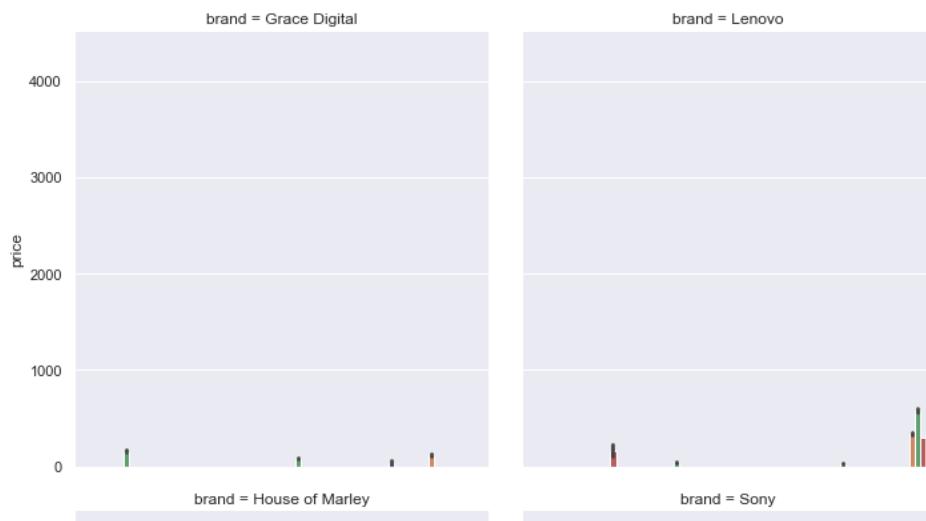
Observation

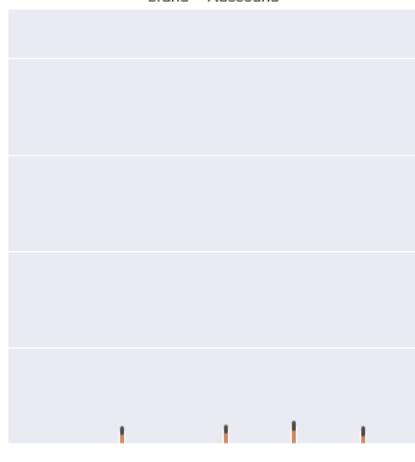
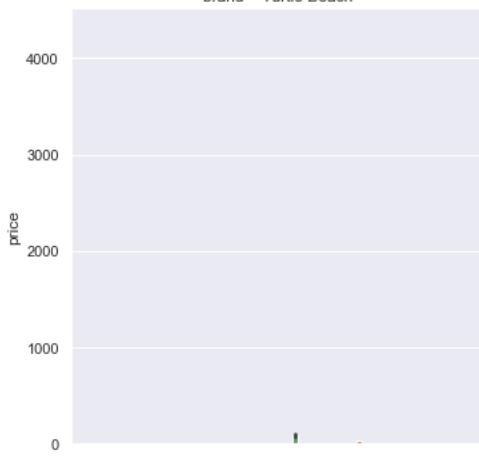
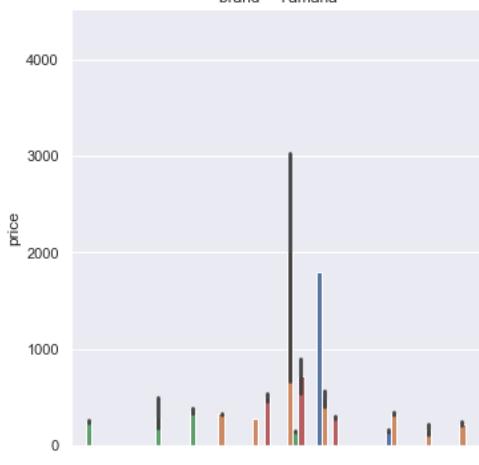
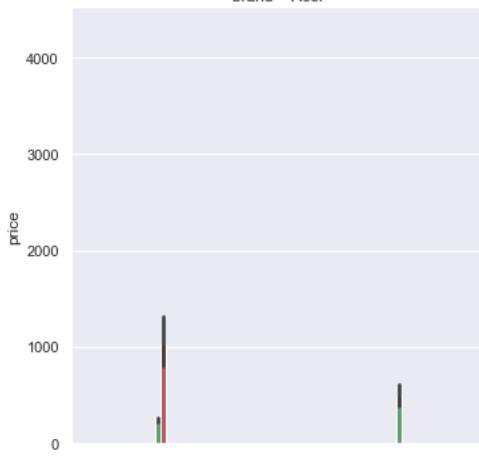
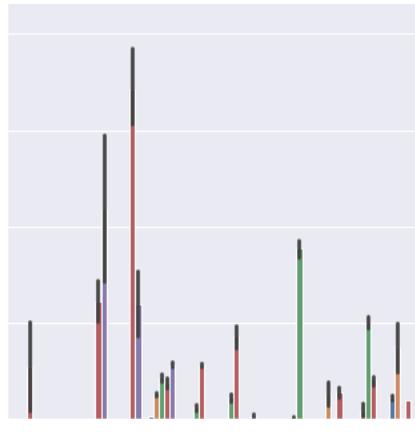
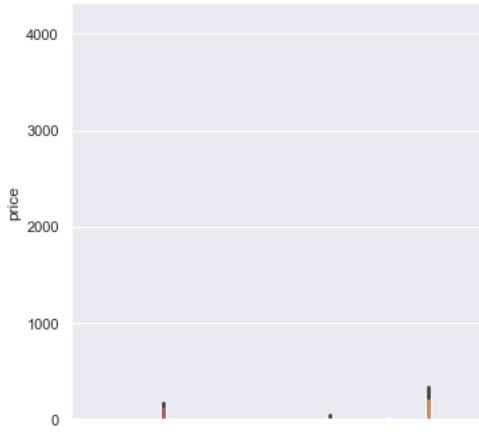
From this graph I can conclude Samsung has more price than other category
I can also conclude that in 3rd month of 2017 price is highest

I will plot brand with dateAdded

In [74]:

```
sns.set(style='darkgrid', font_scale=1.0)
sns.catplot(x='Month_dtadd',
            y='price', hue='Year_dtadd',
            col='brand',
            col_wrap=2,
            data=dt_electronics,
            kind='bar')
plt.savefig('Dateadded_brand.png')
```



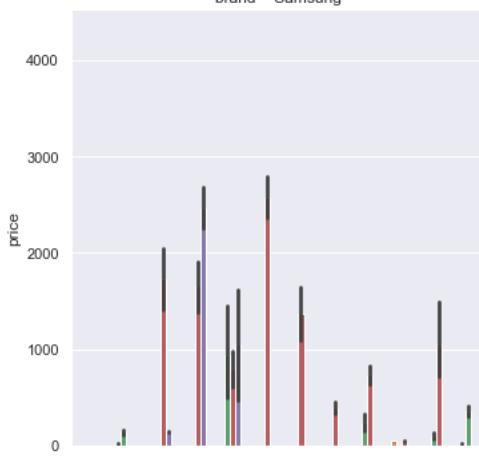




brand = Samsung



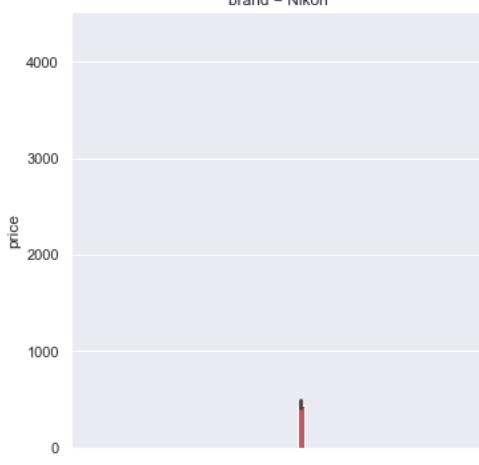
brand = WD



brand = Nikon



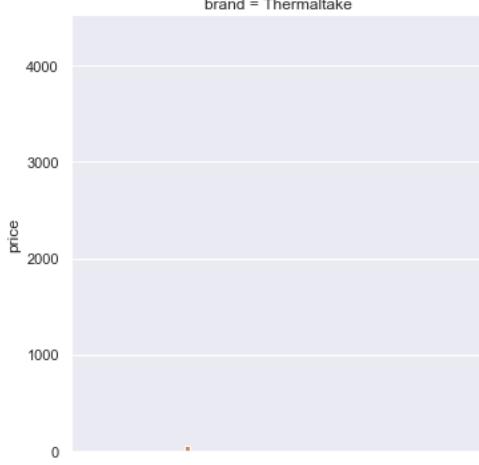
brand = ASUS



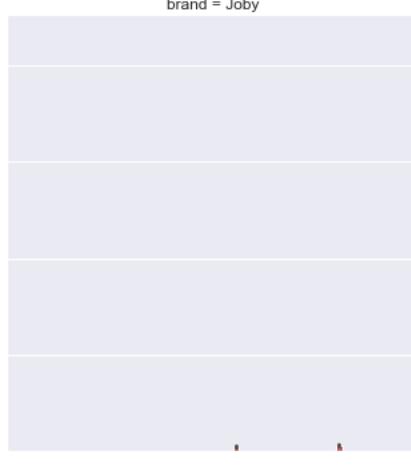
brand = Thermaltake



brand = Joby



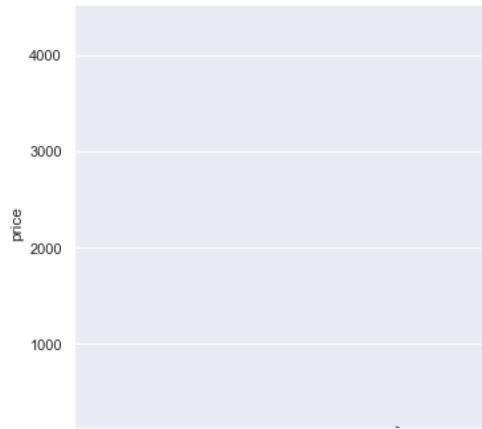
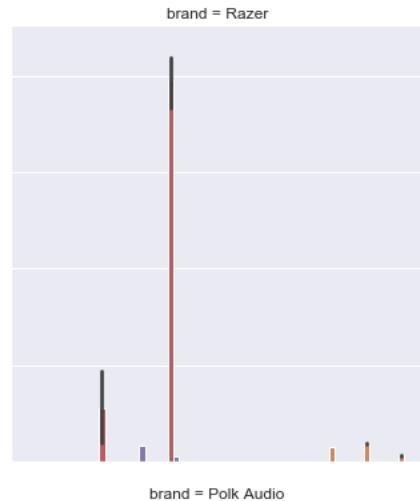
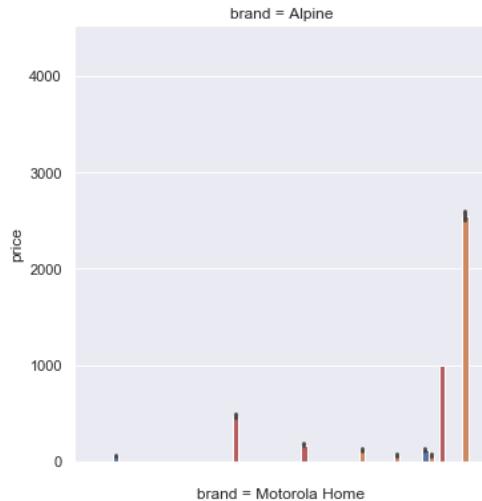
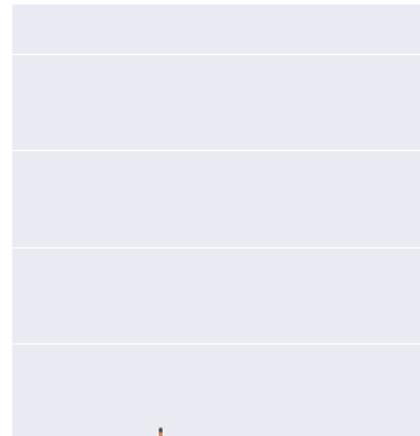
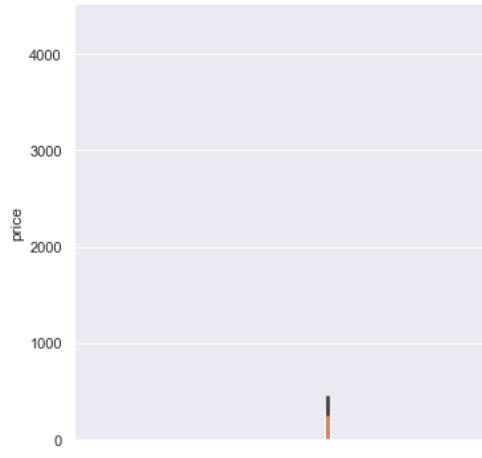
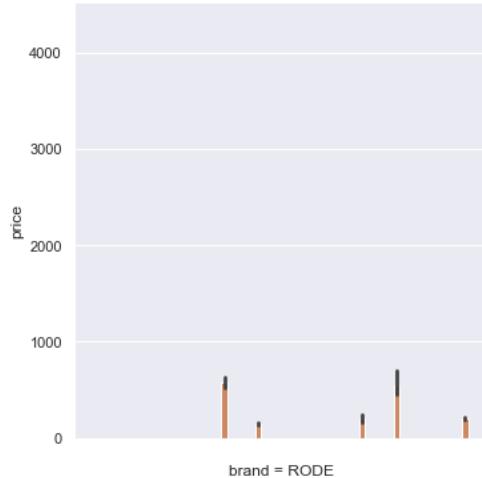
brand = Power Acoustik

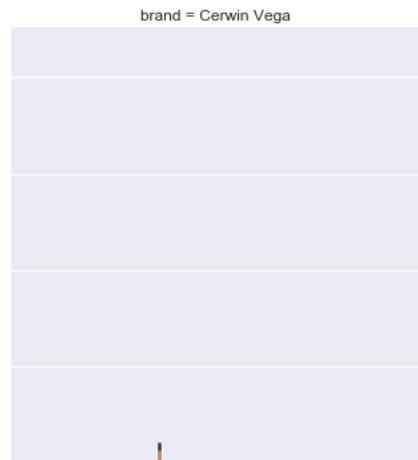
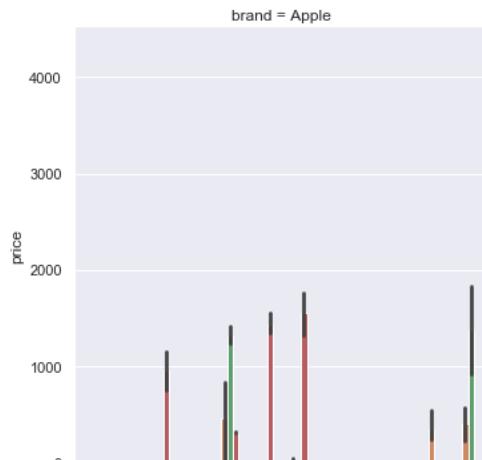
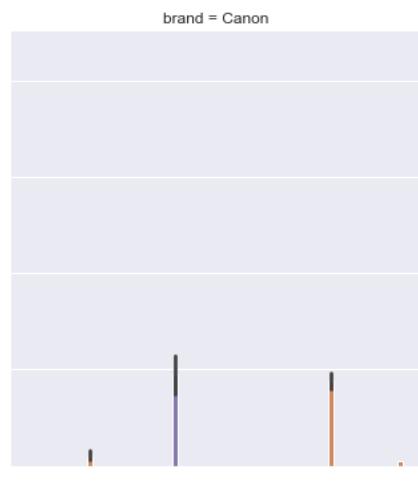
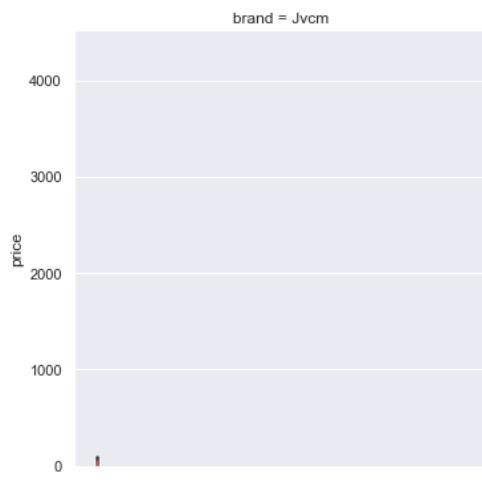
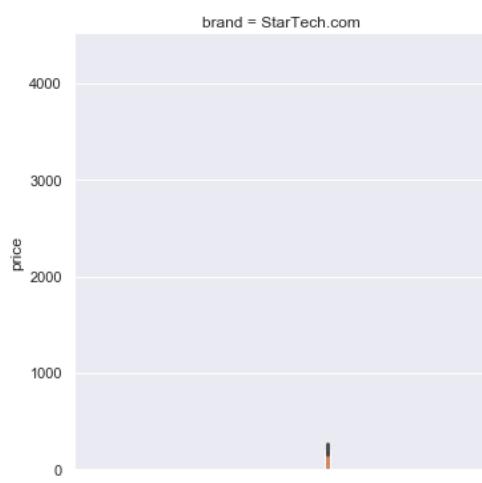
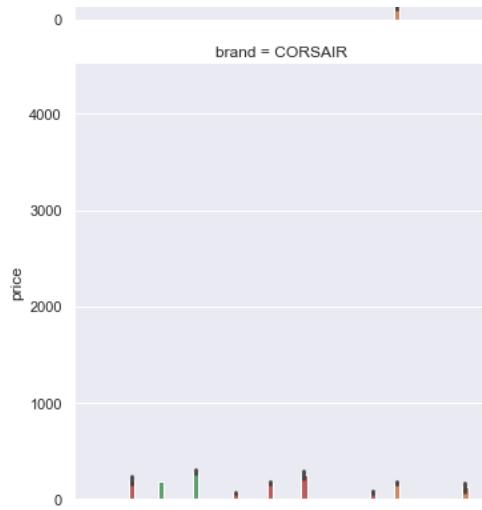


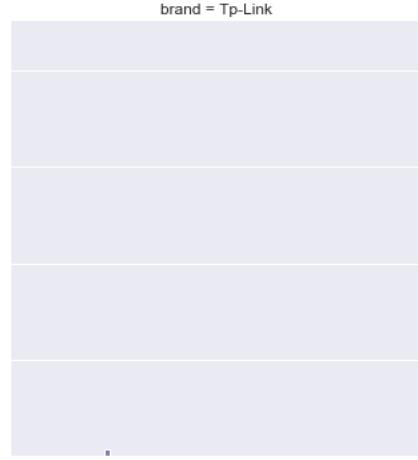
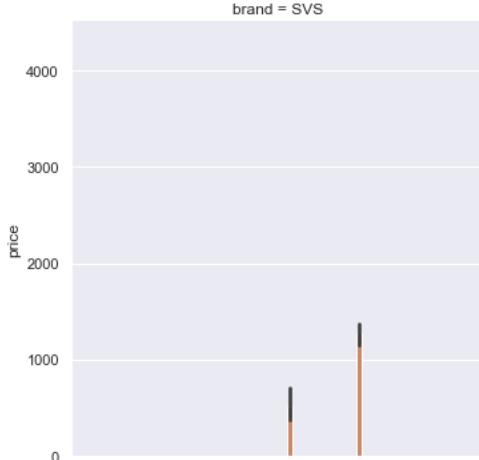
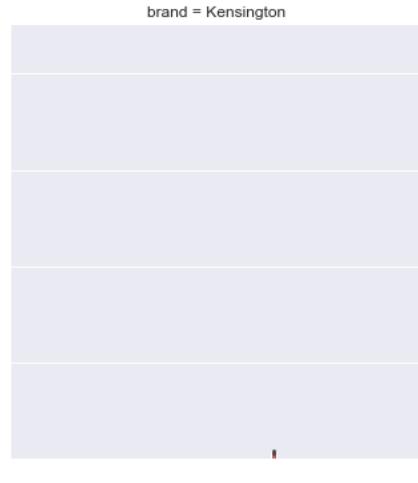
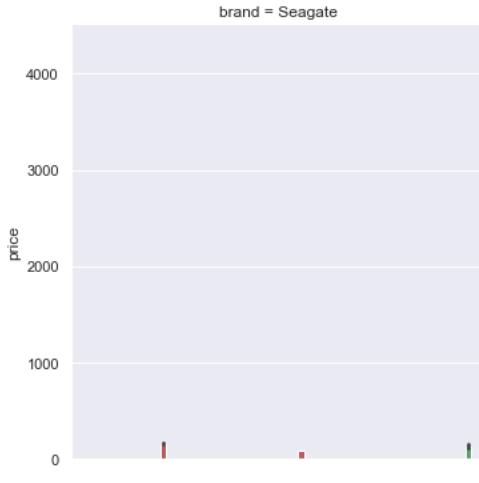
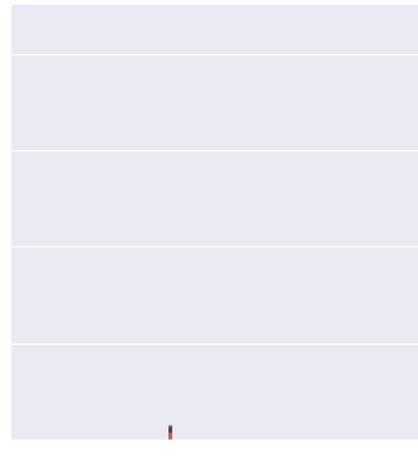
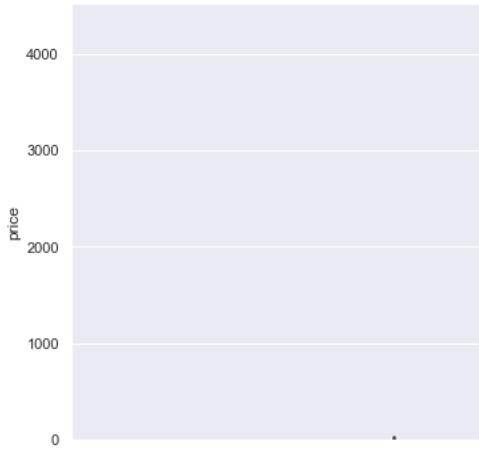
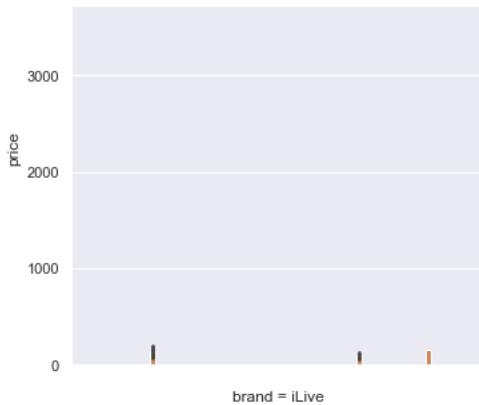
brand = Sonic Alert

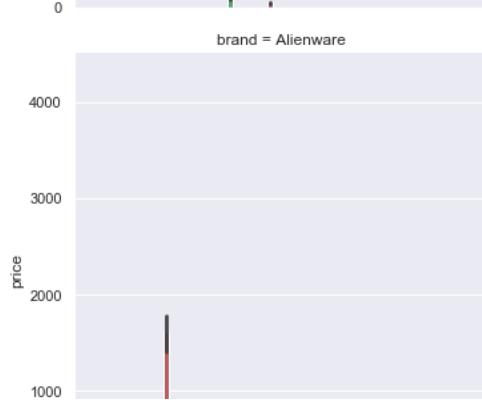
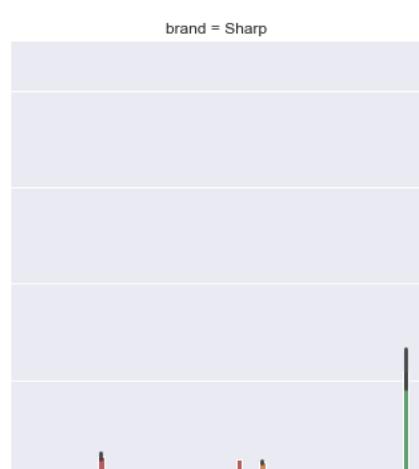
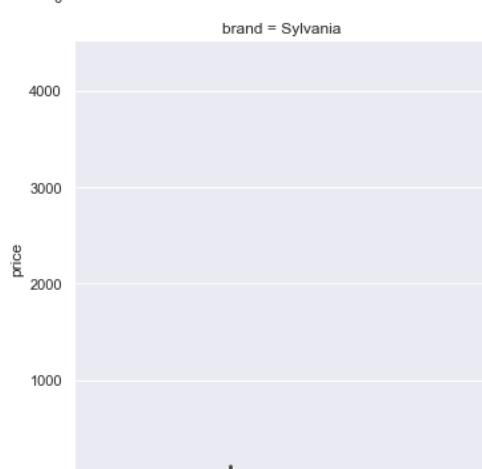
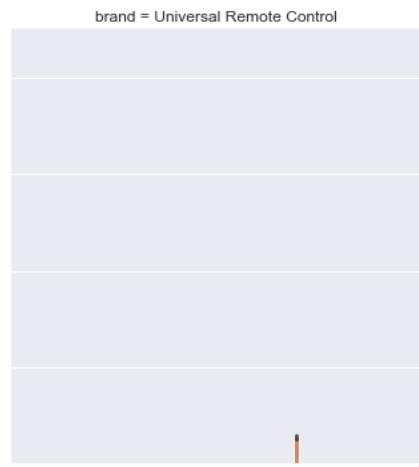
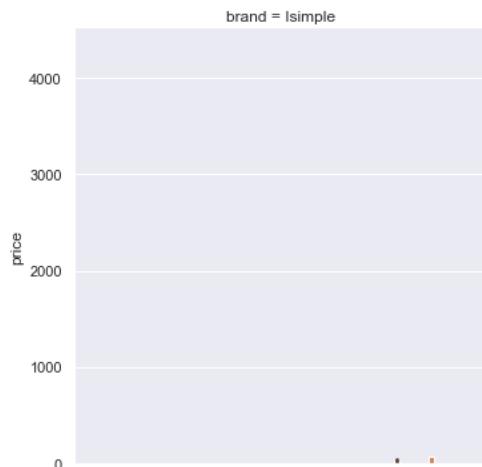
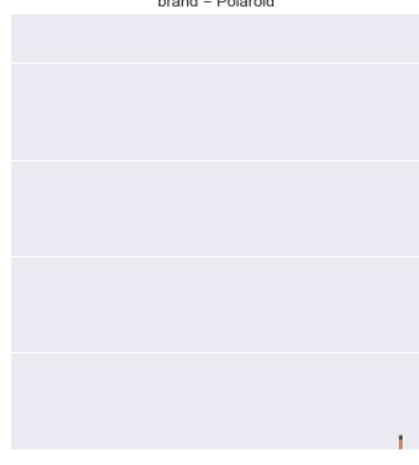
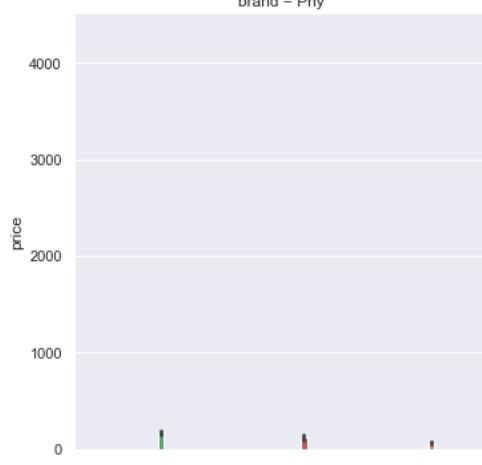


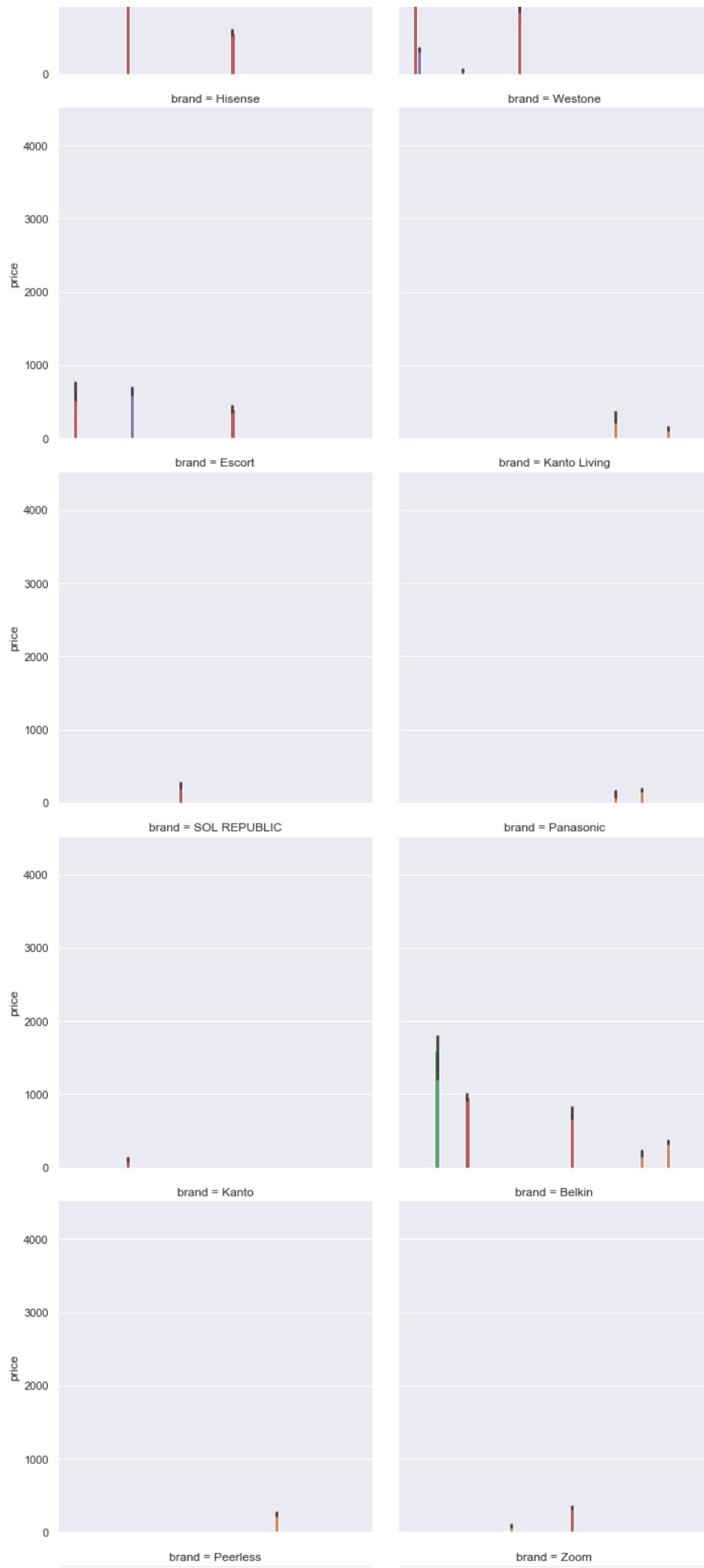
price

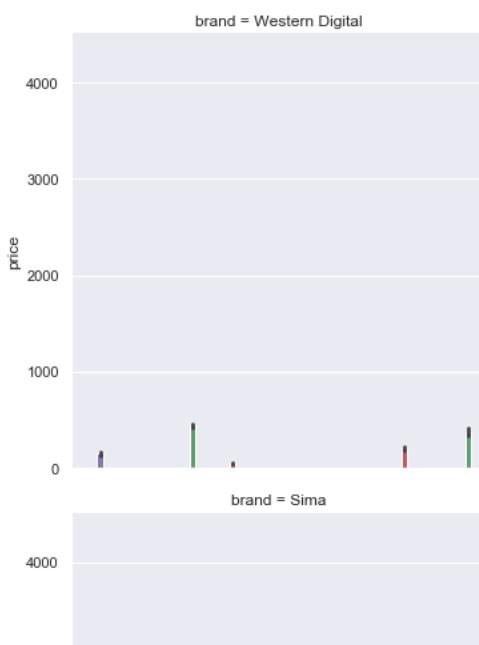
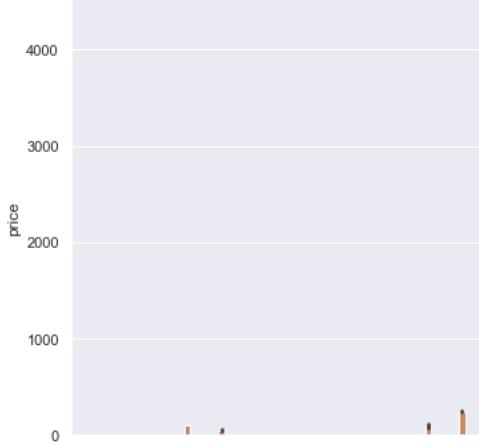
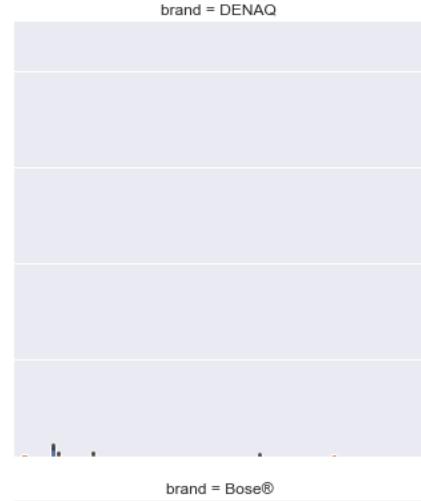
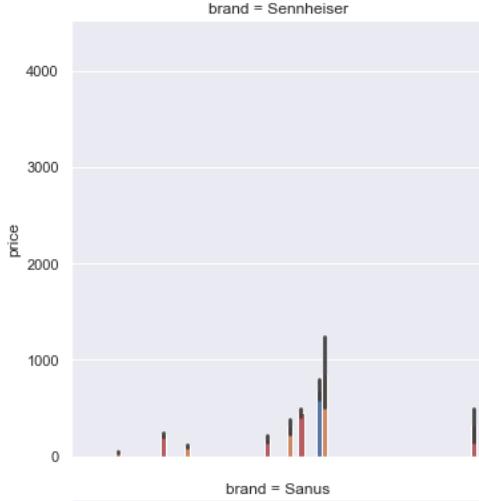
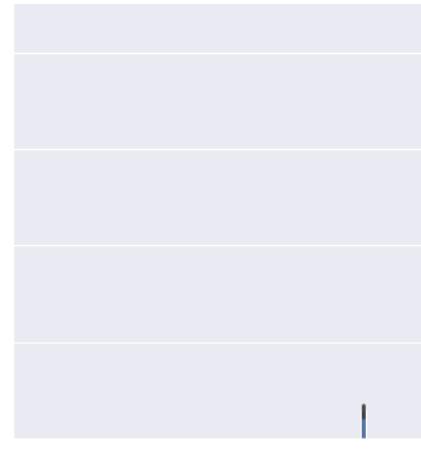
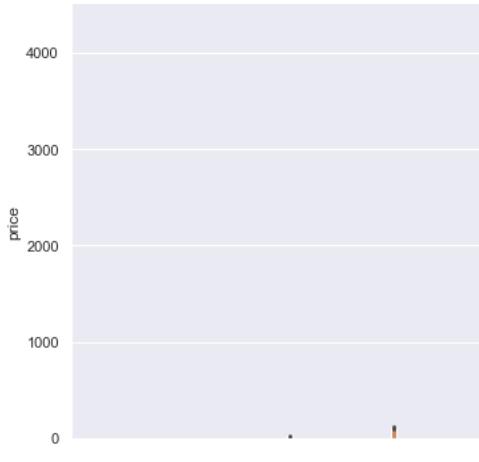


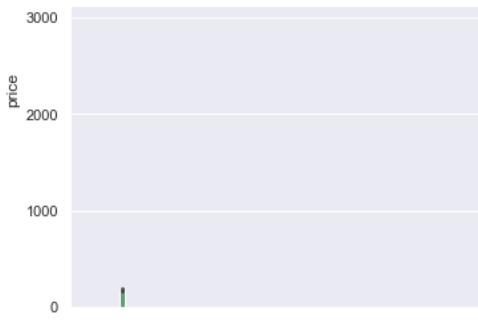












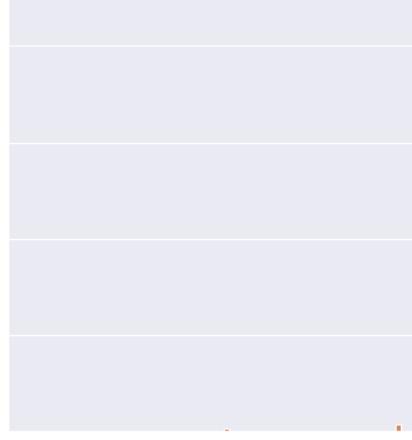
brand = V-MODA



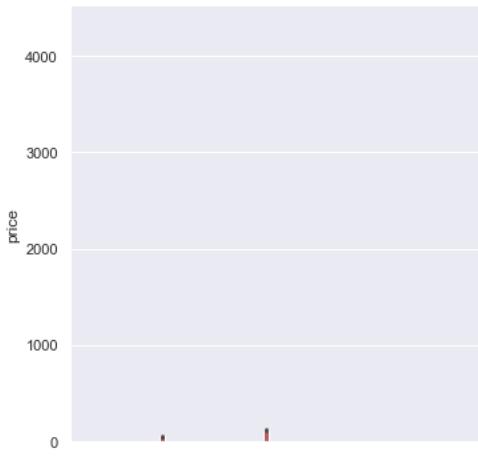
brand = Manfrotto



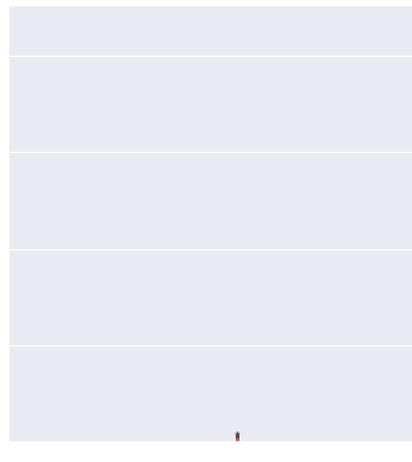
brand = Monster



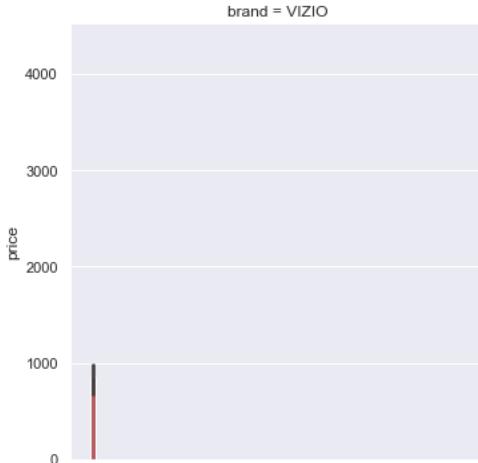
brand = Zmodo



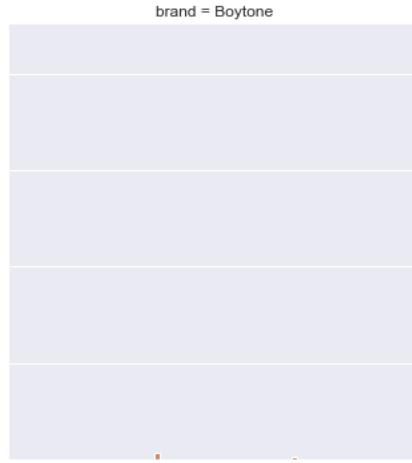
brand = VIZIO



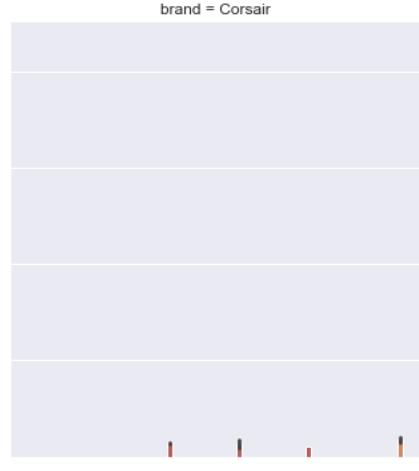
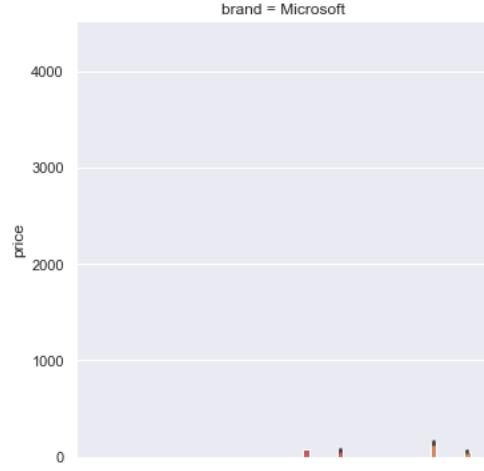
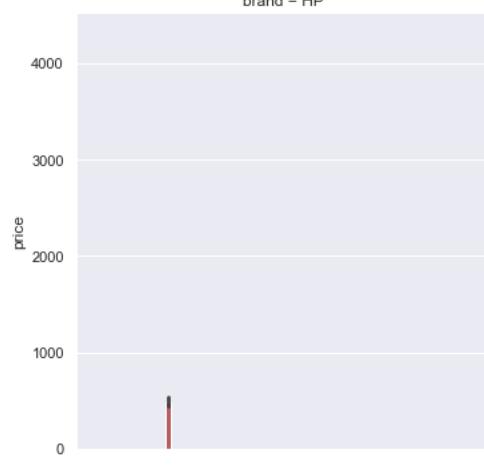
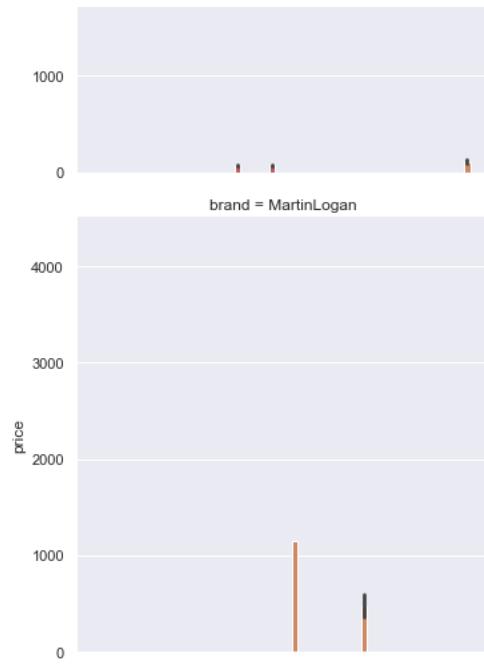
brand = Boytone

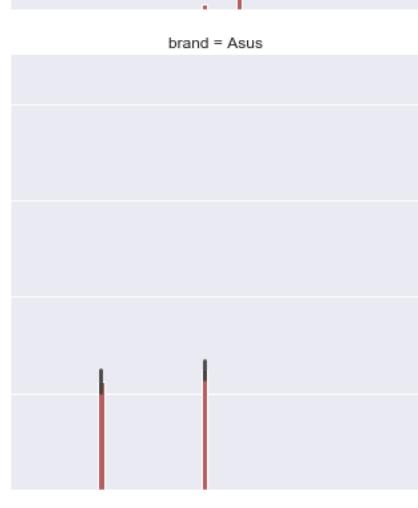
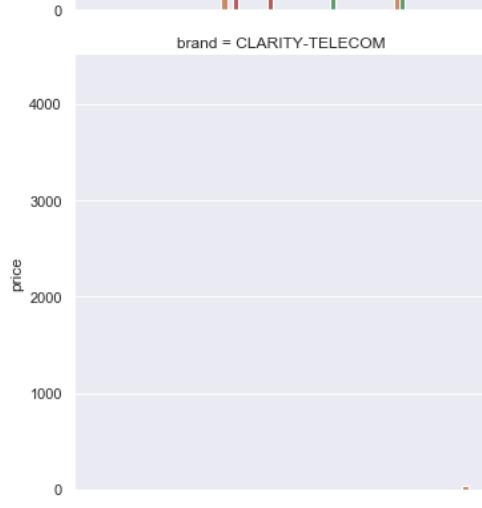
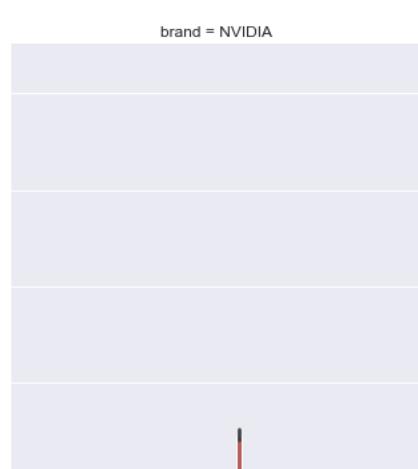
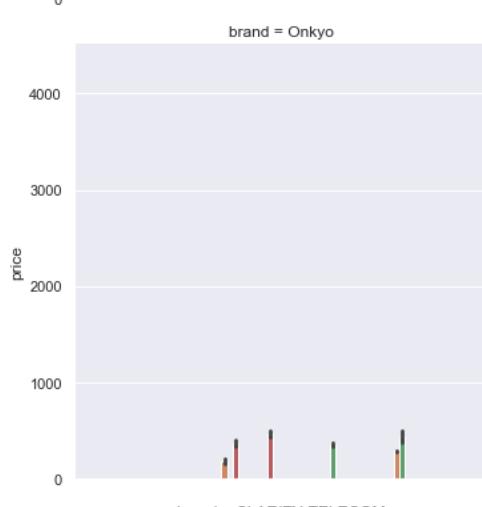
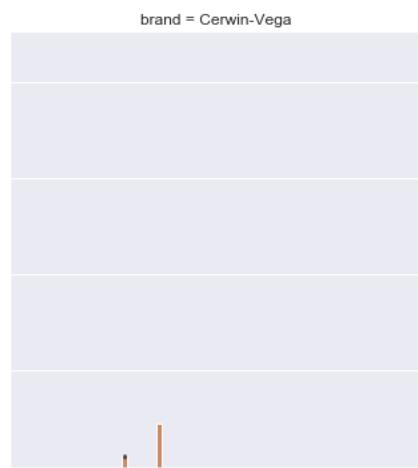
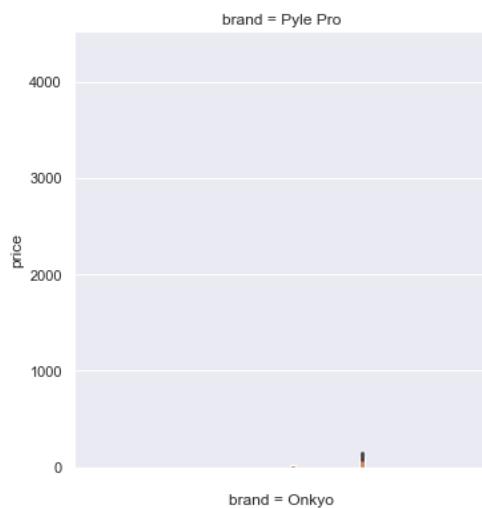
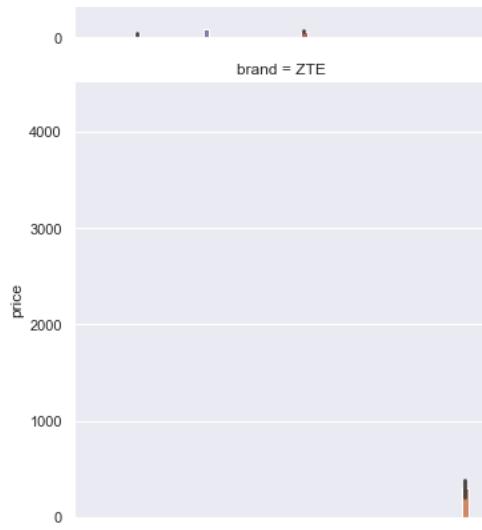


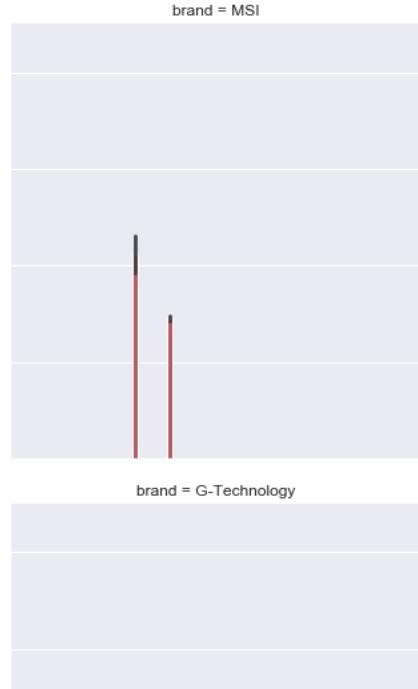
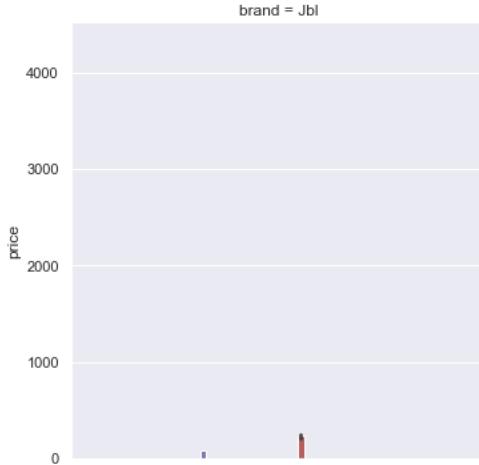
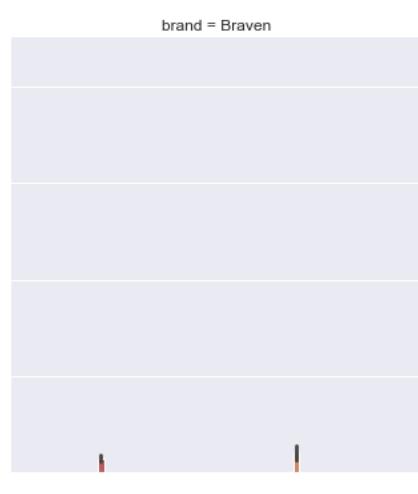
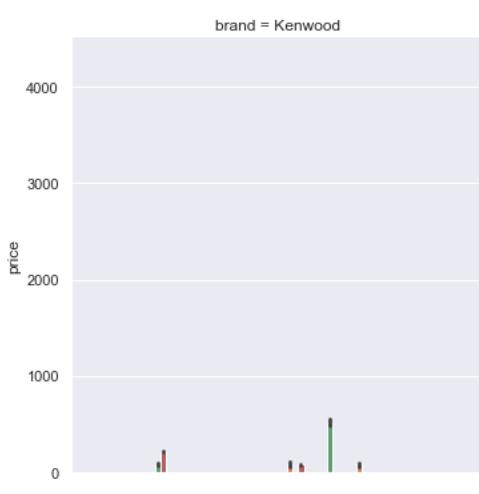
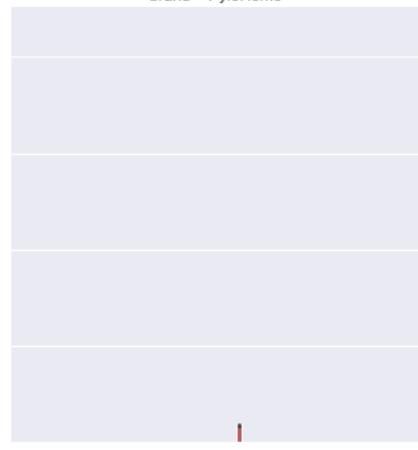
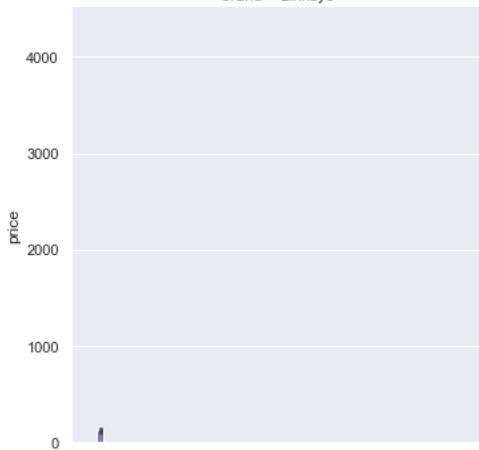
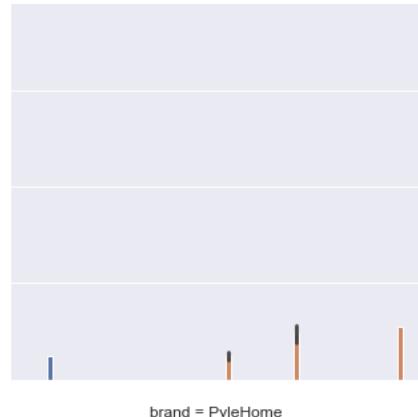
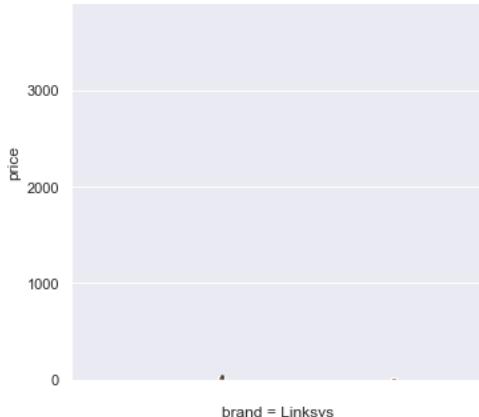
brand = Siriusxm



brand = ECOXGEAR

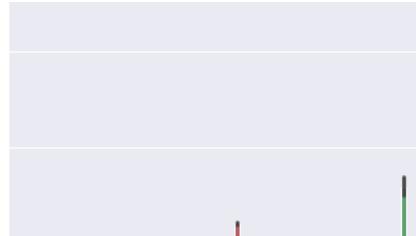




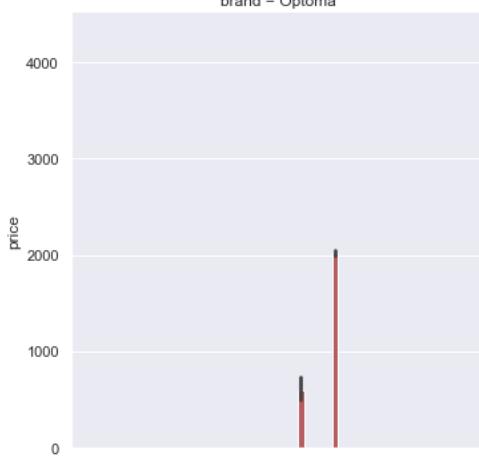




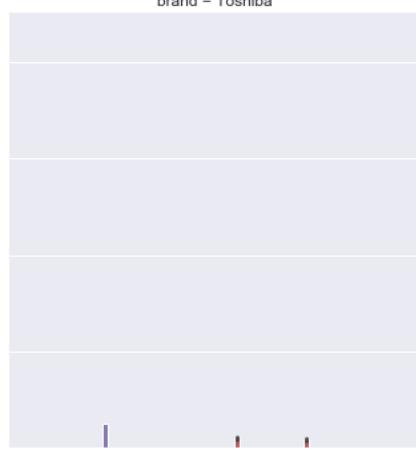
brand = Optoma



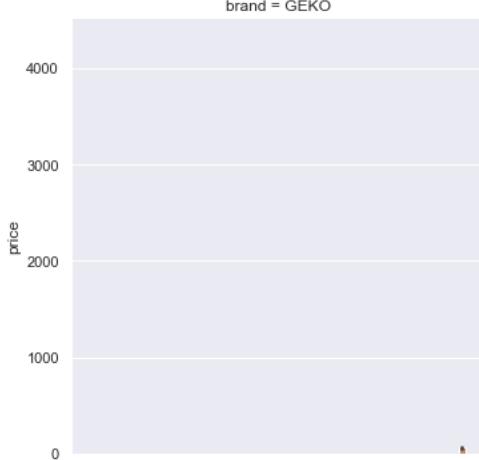
brand = Toshiba



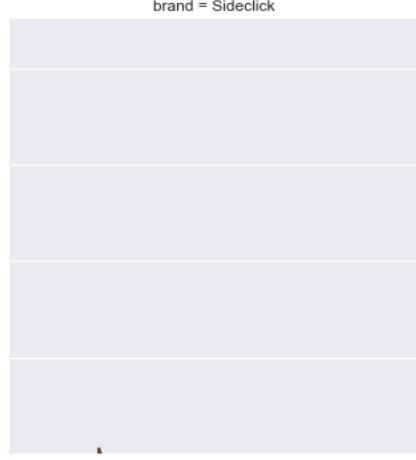
brand = GEKO



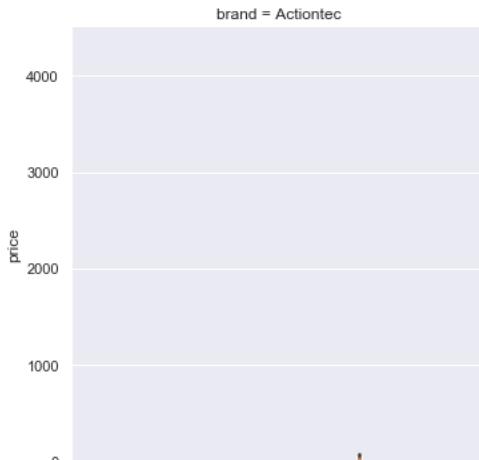
brand = Sideclick



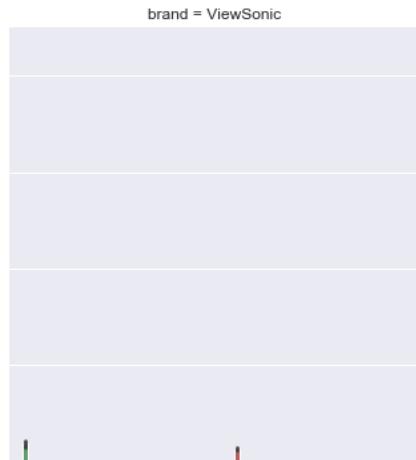
brand = Actiontec



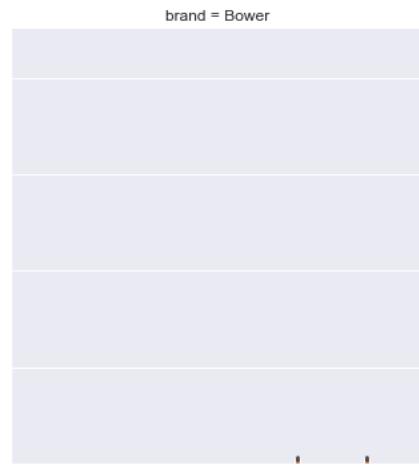
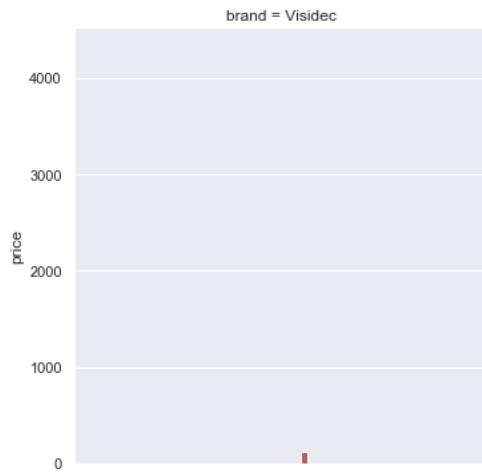
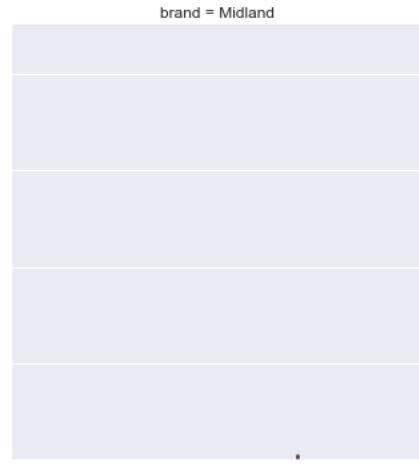
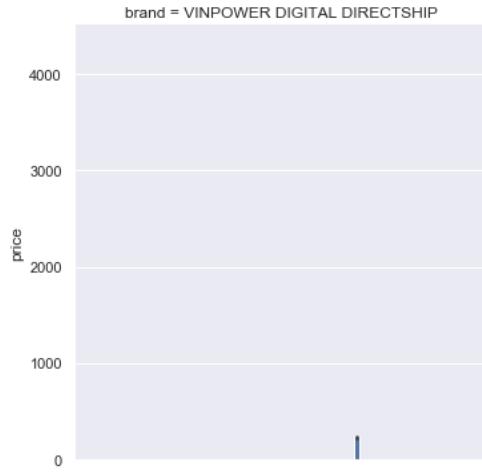
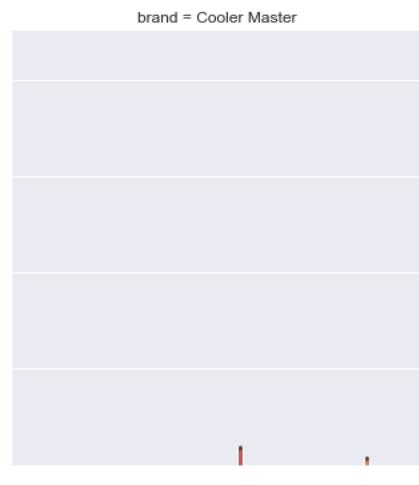
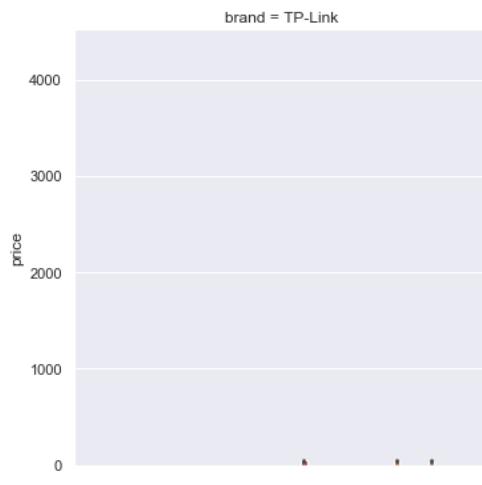
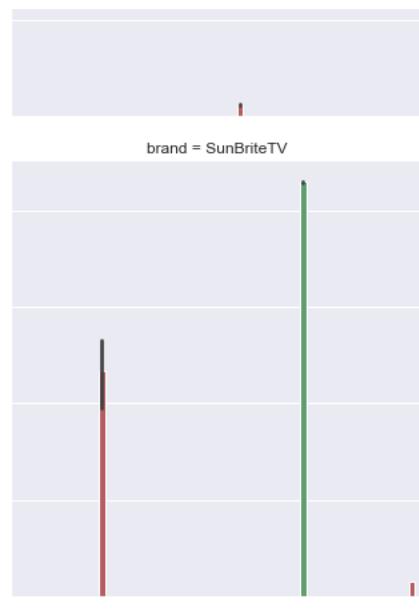
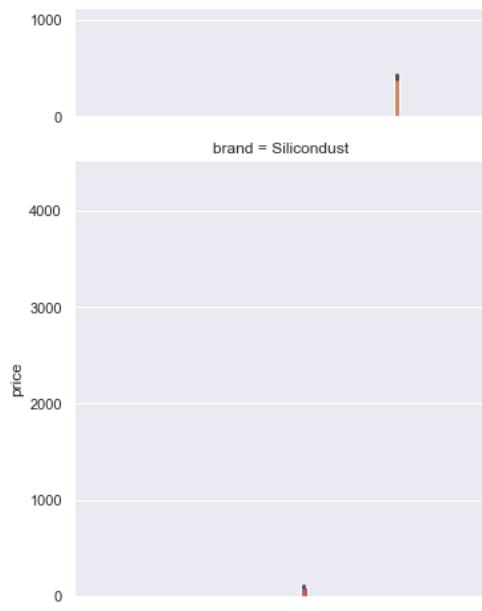
brand = ViewSonic

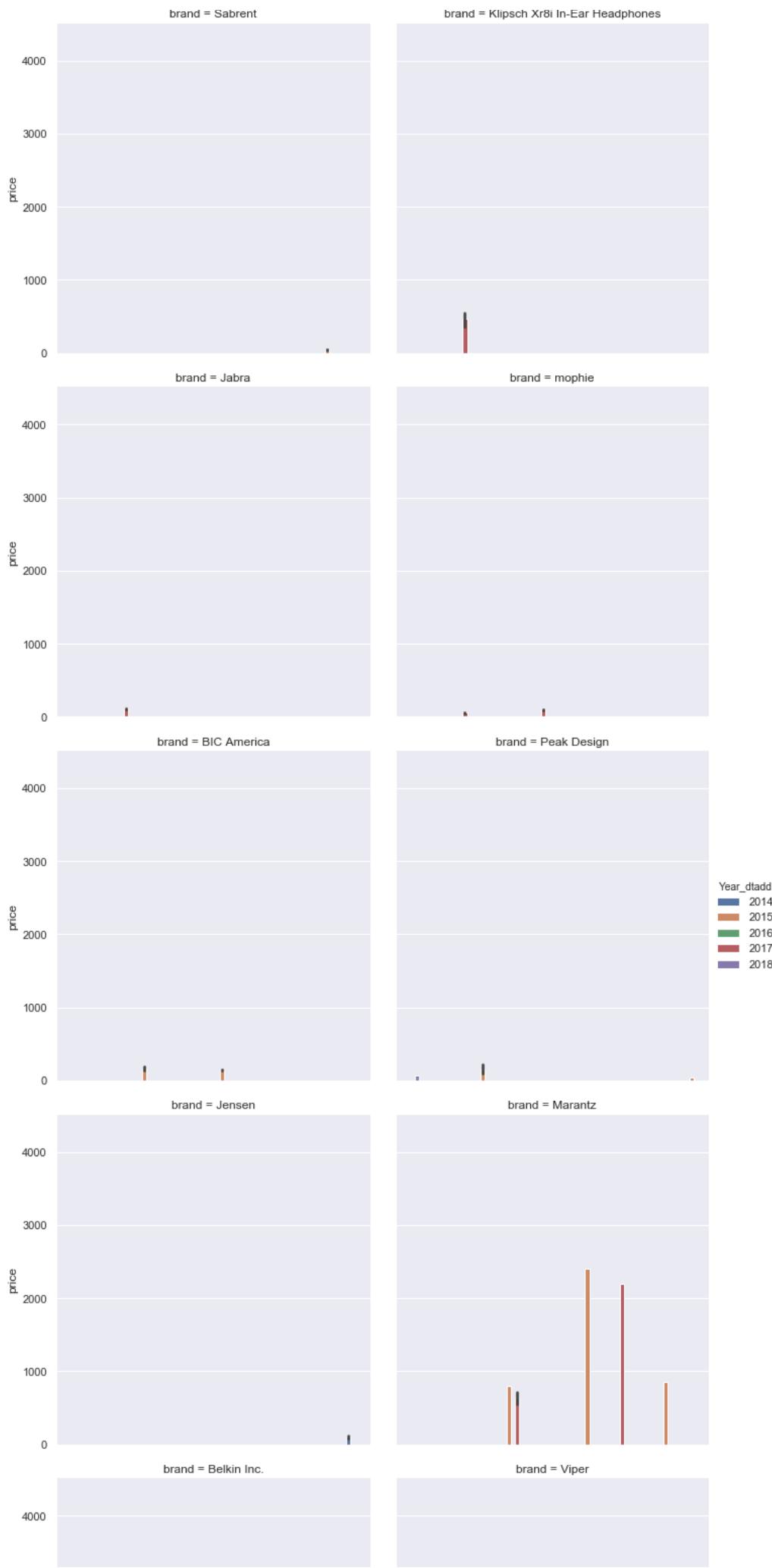


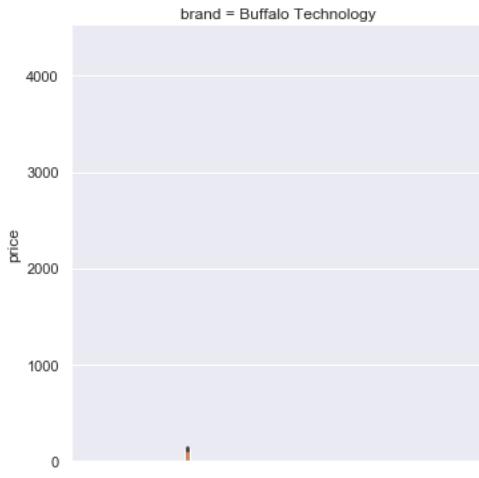
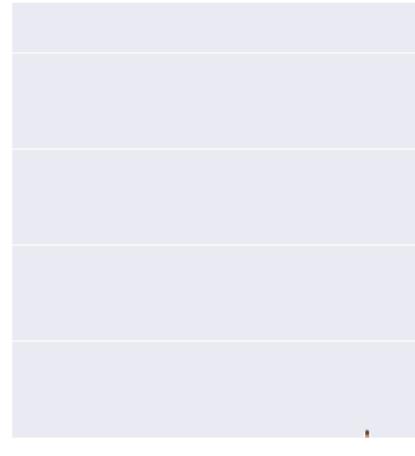
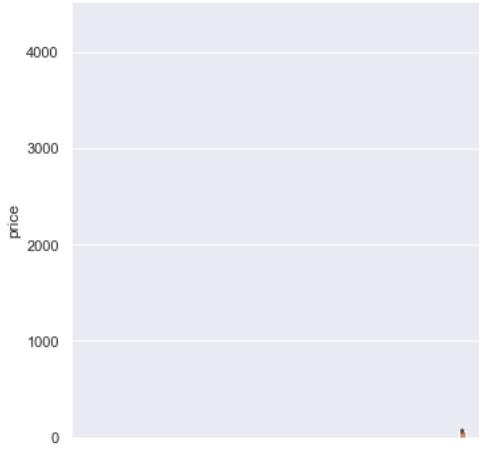
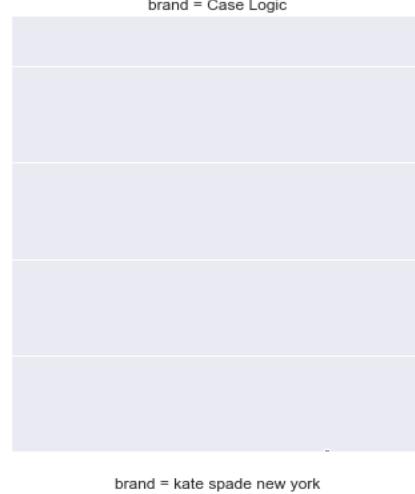
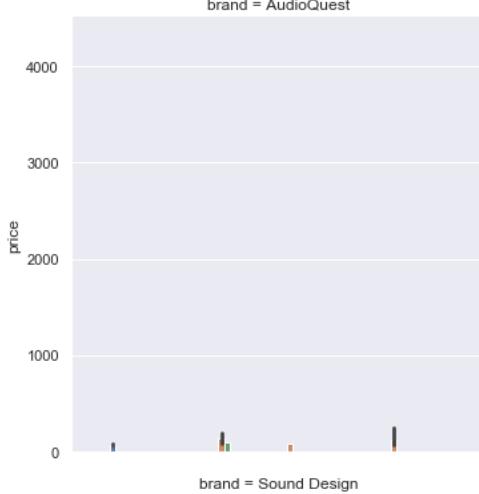
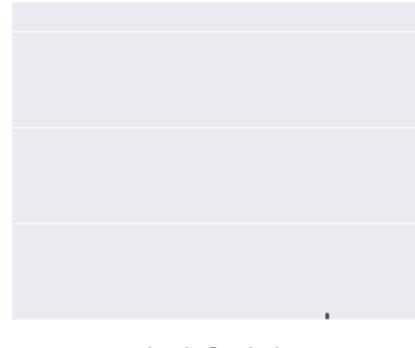
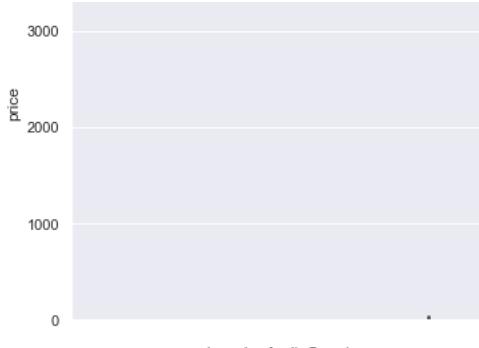
brand = Buffalo

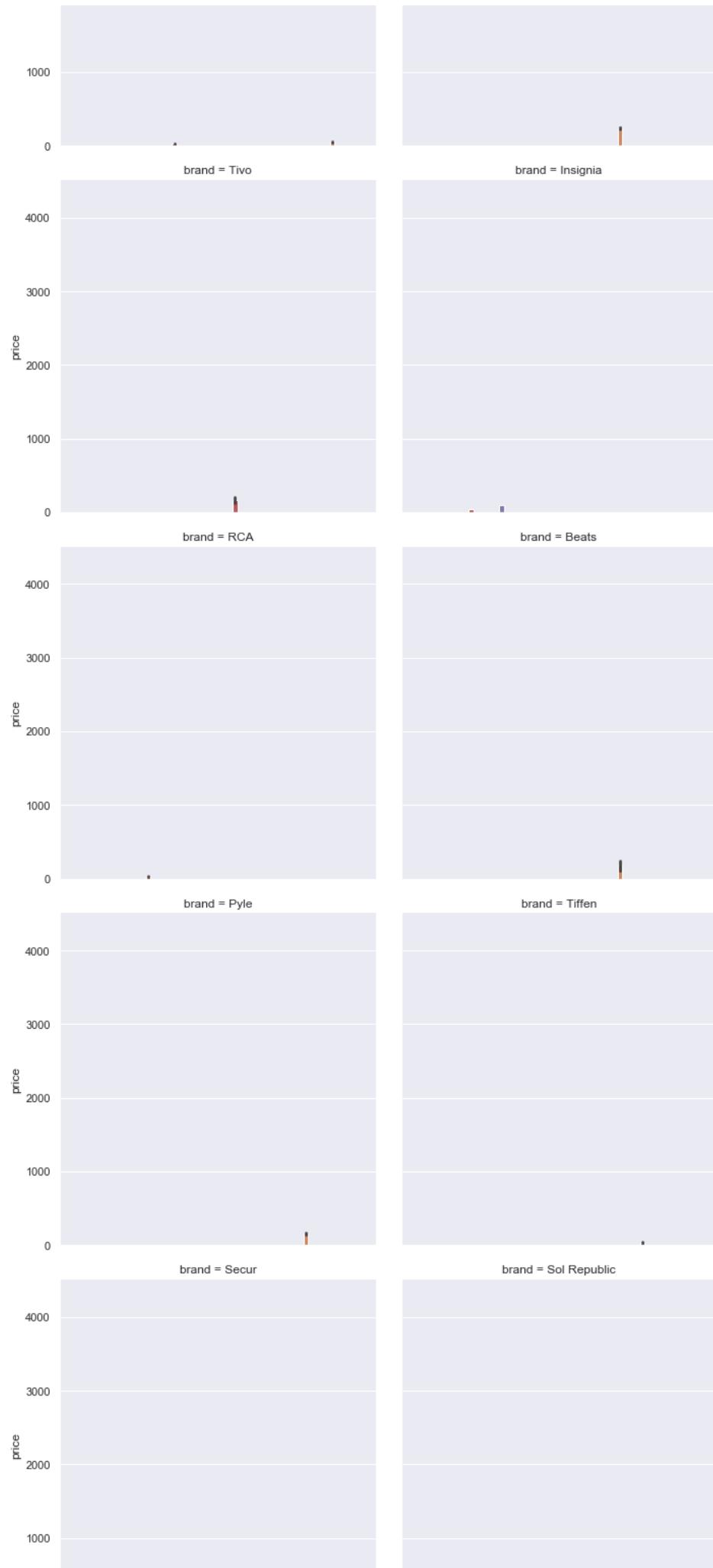


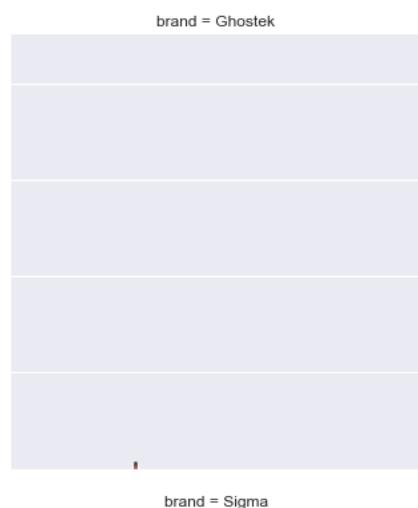
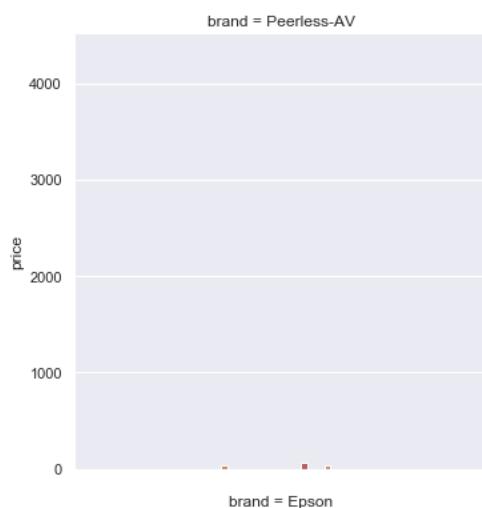
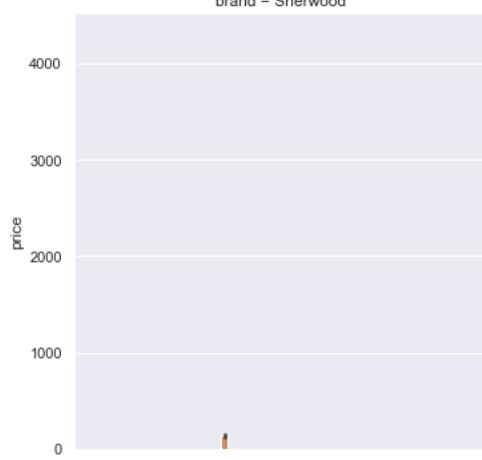
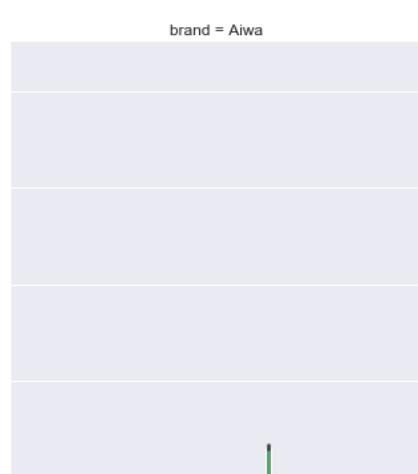
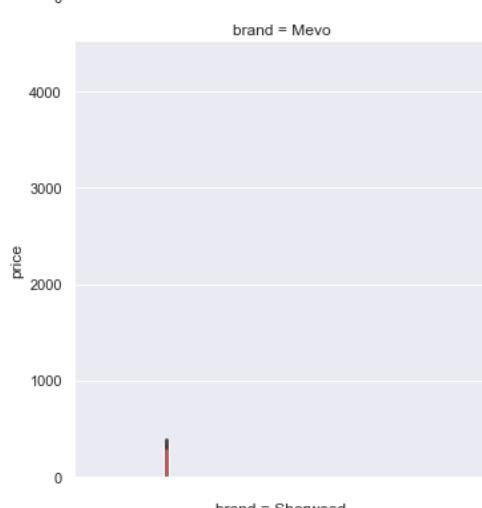
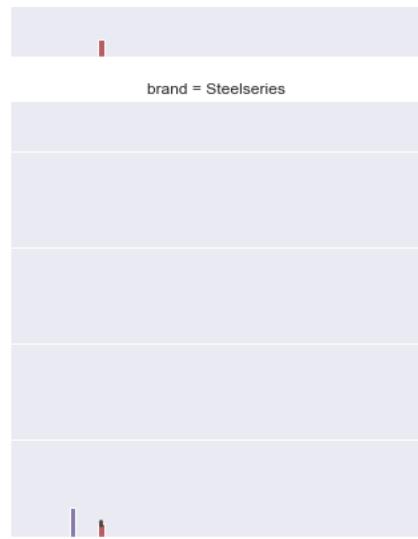
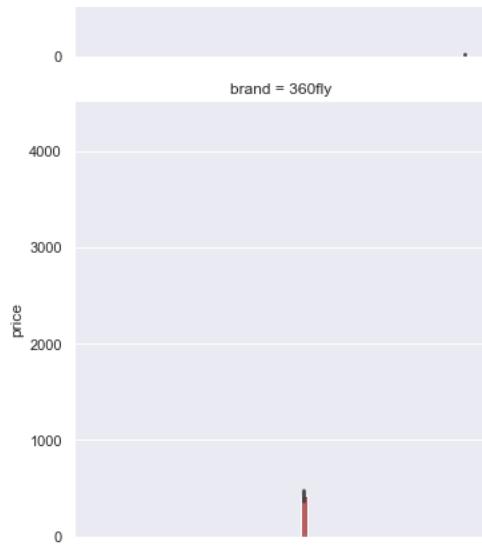
brand = Cobra Electronics









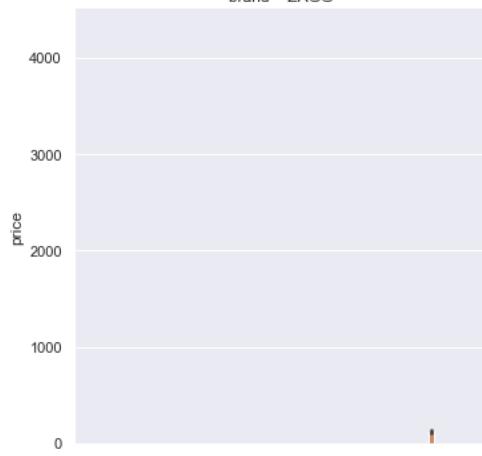




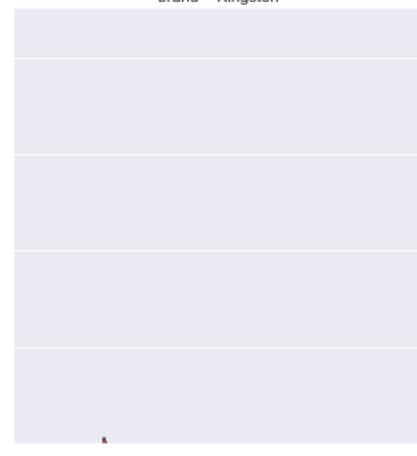
brand = ZAGG



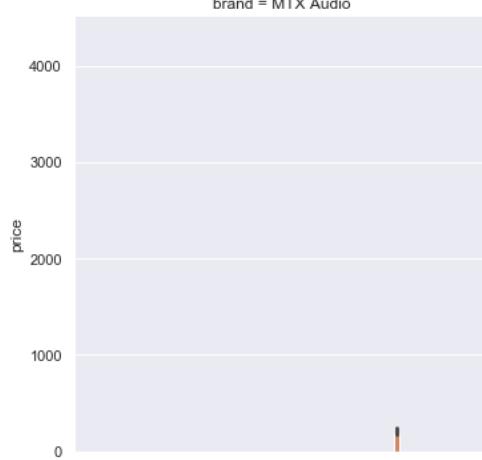
brand = Kingston



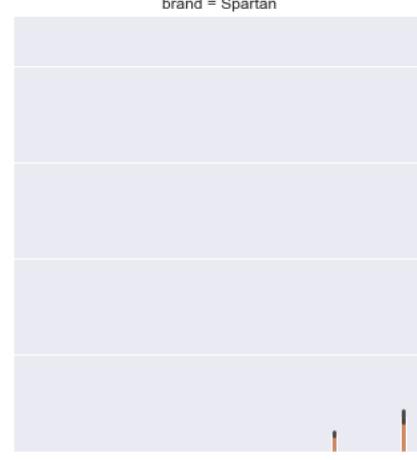
brand = MTX Audio



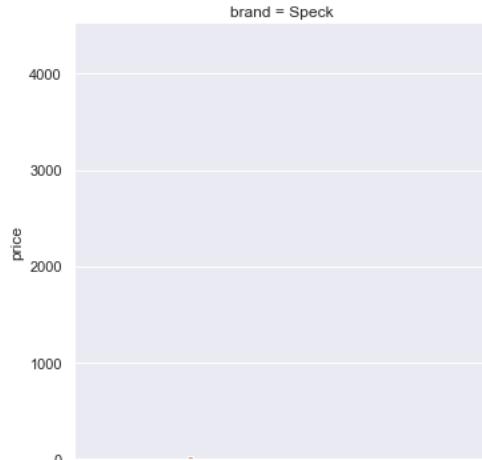
brand = Spartan



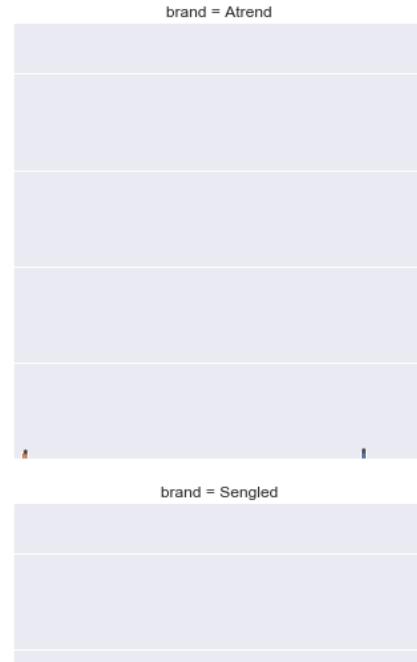
brand = Speck



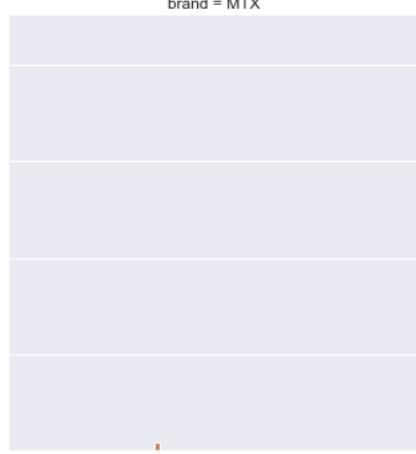
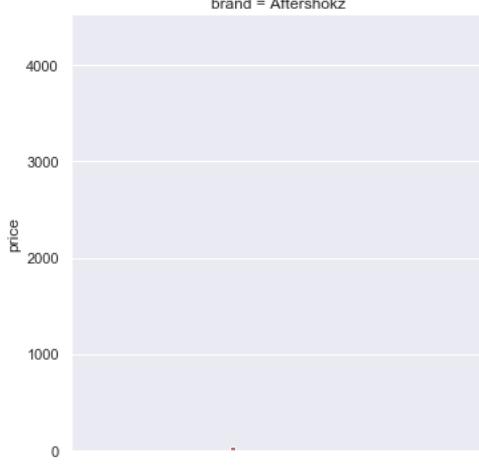
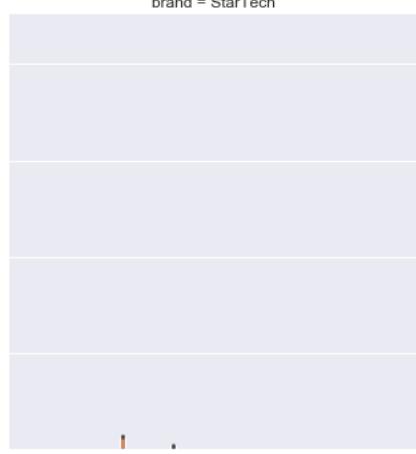
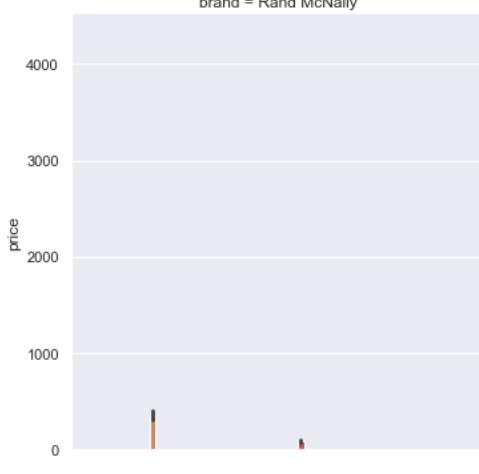
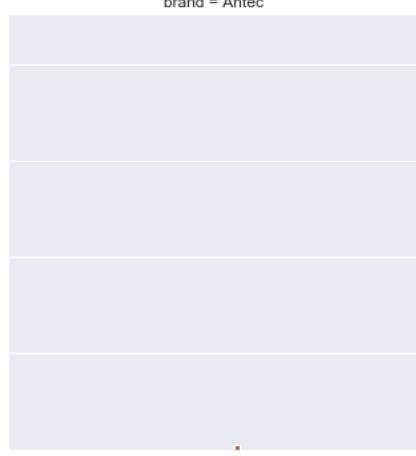
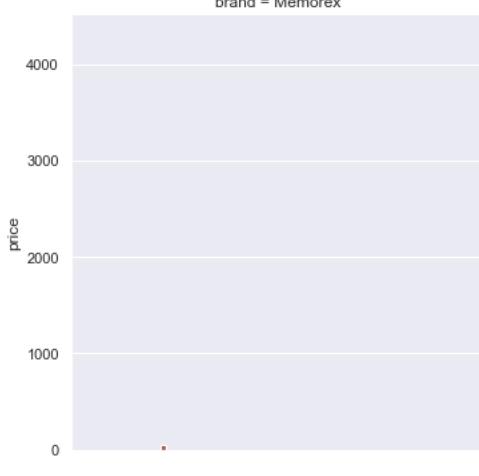
brand = Atrend

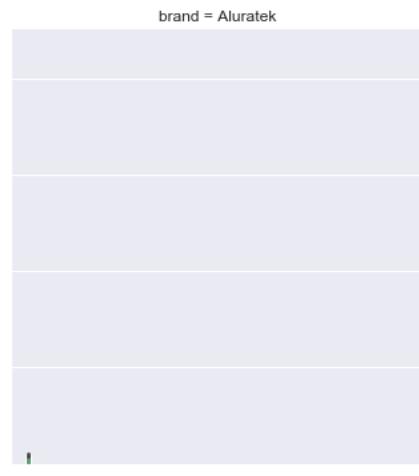
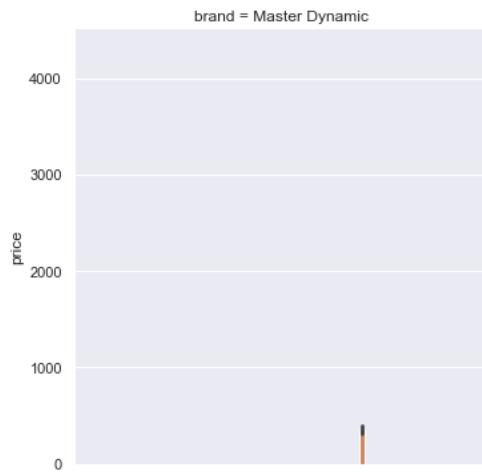
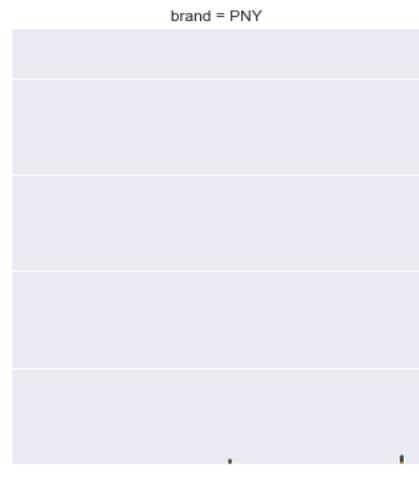
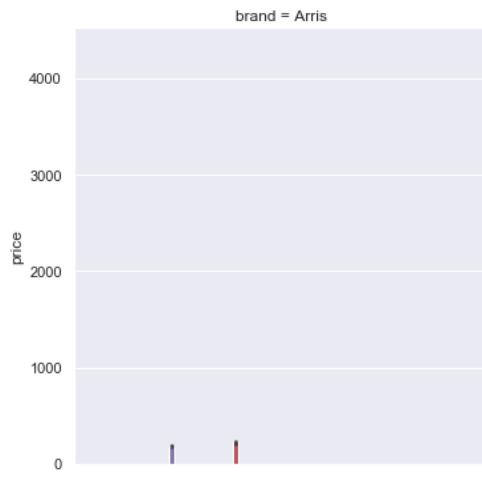
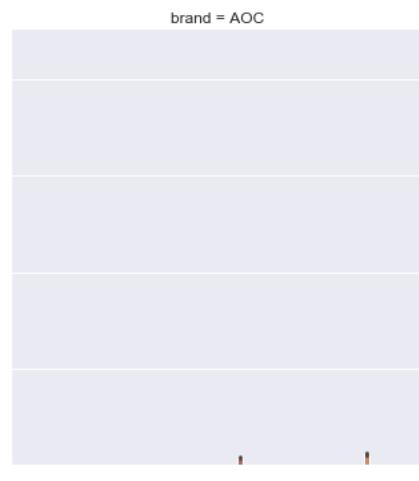
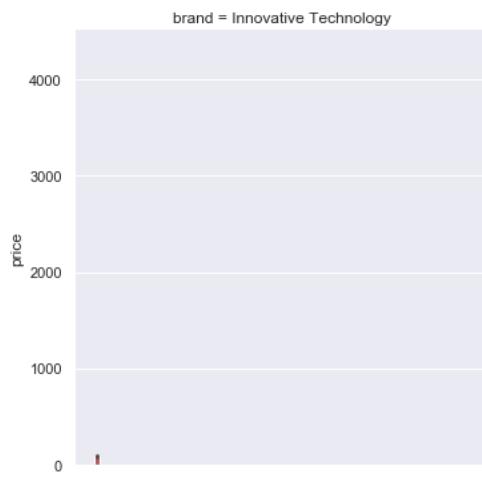
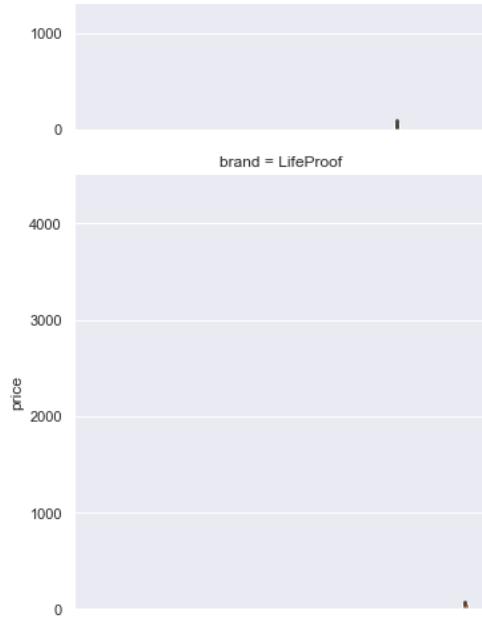


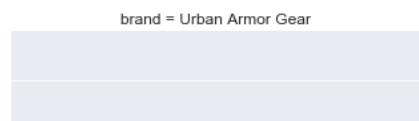
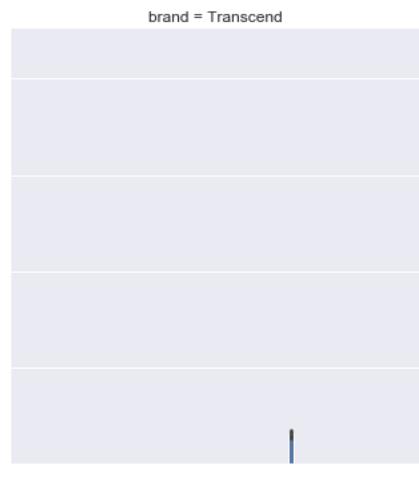
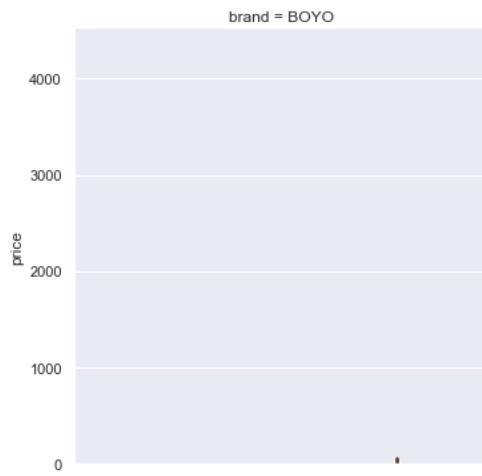
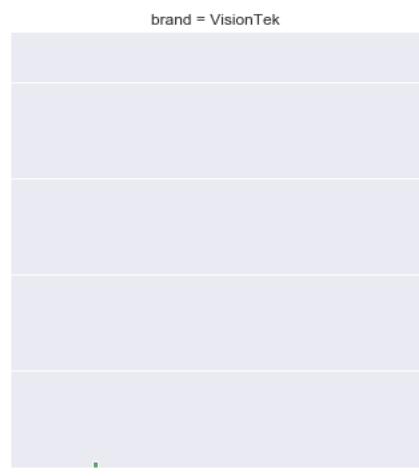
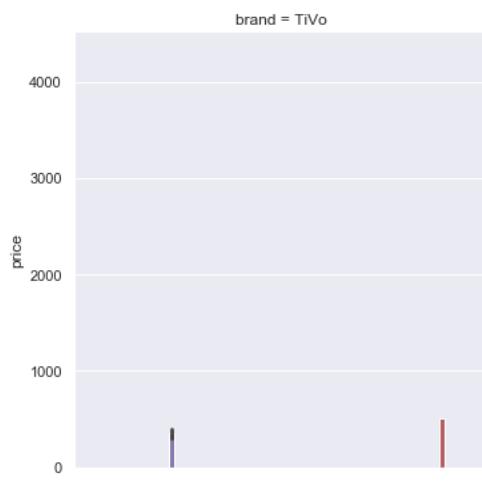
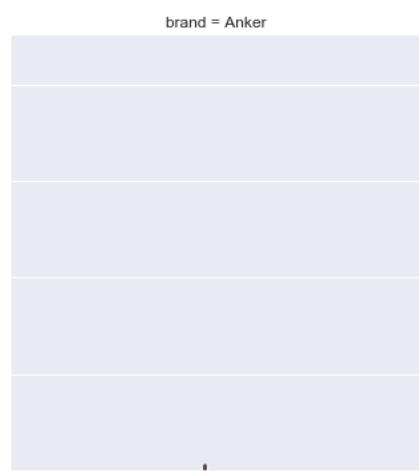
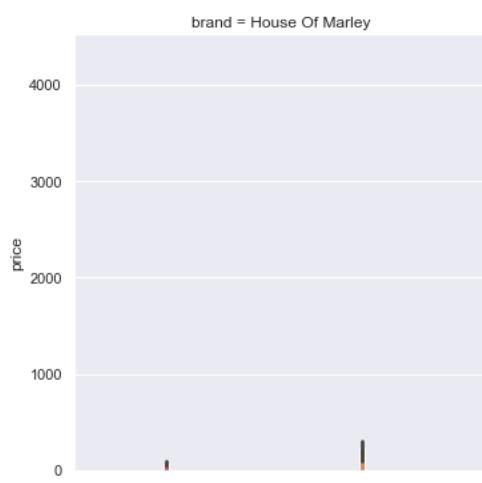
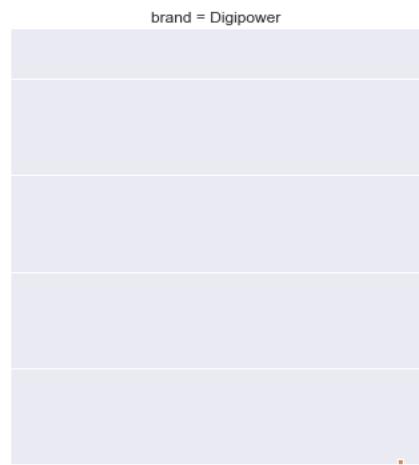
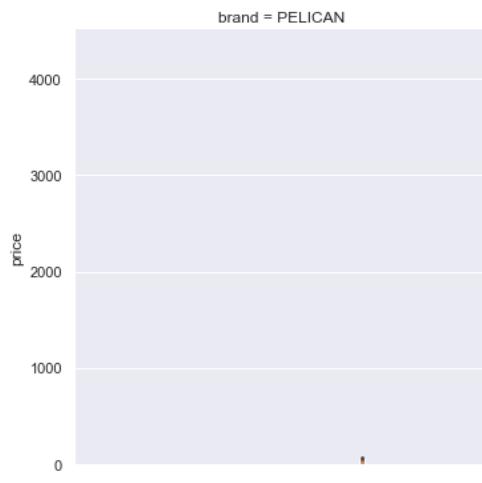
brand = TCL

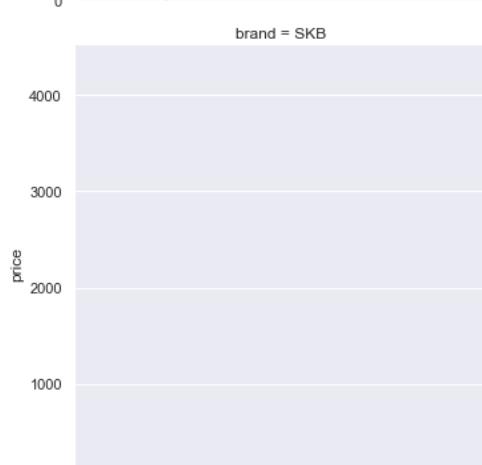
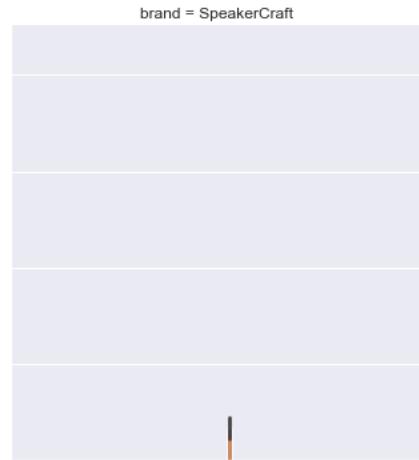
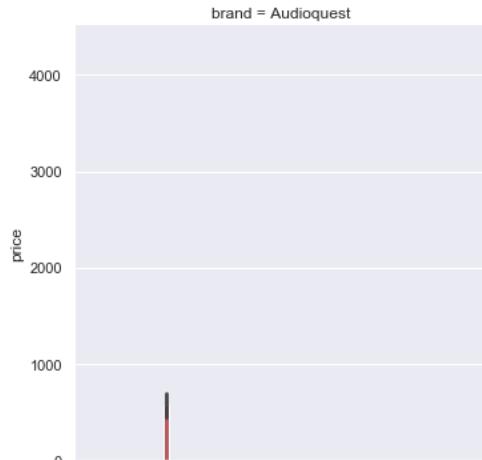
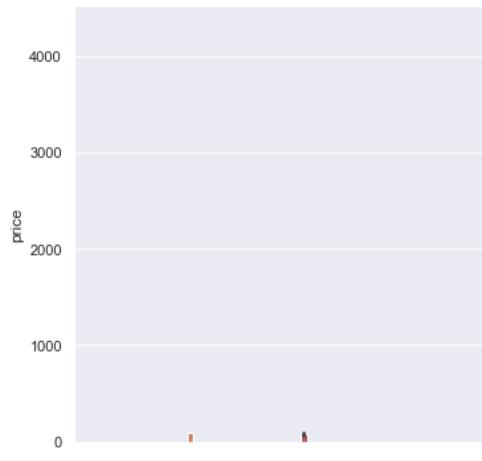
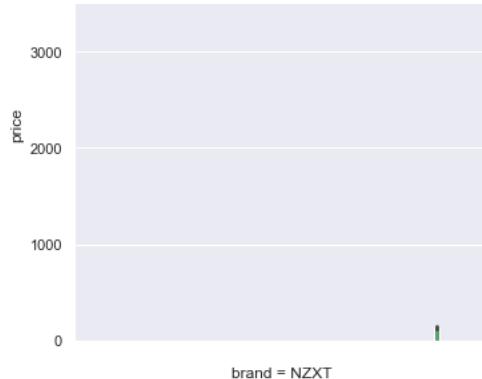


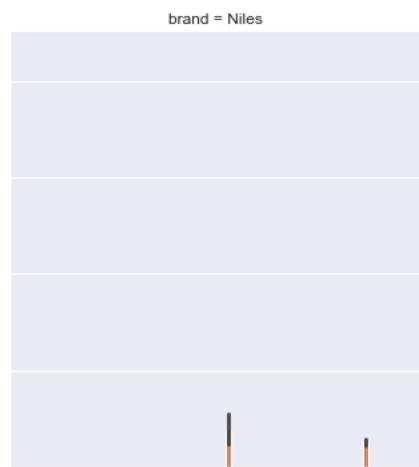
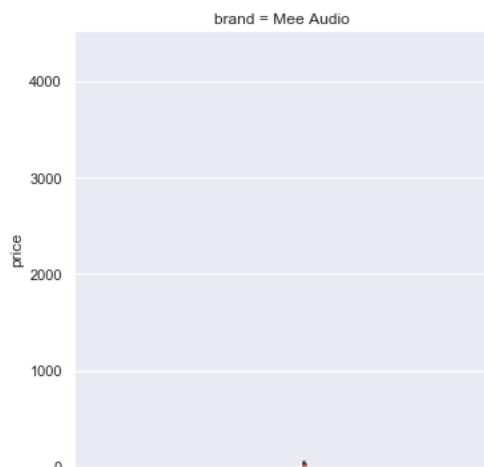
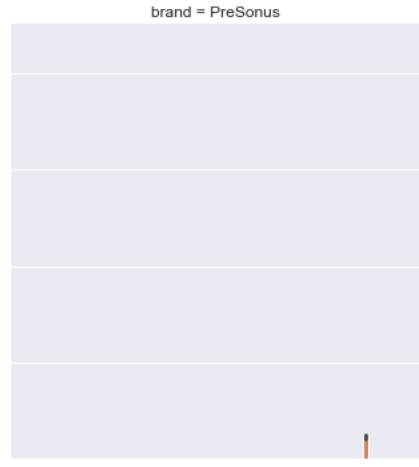
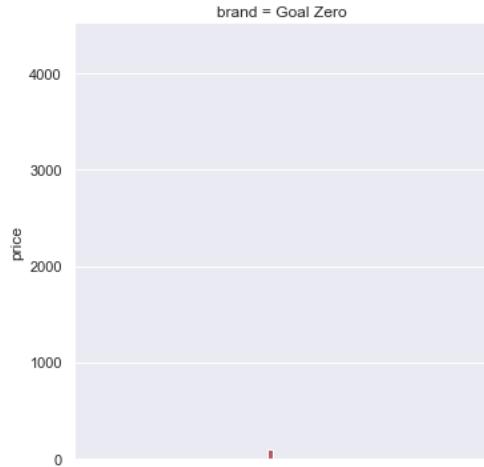
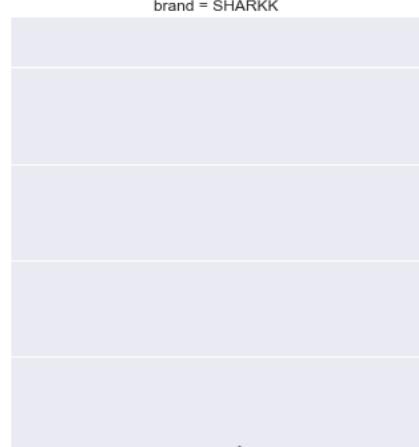
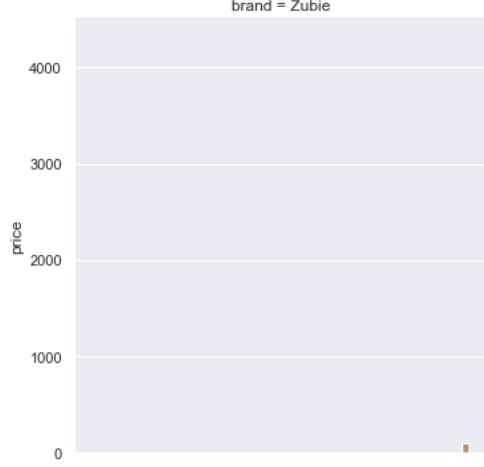
brand = Sengled

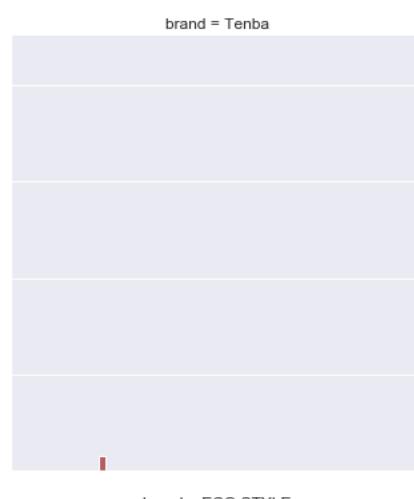
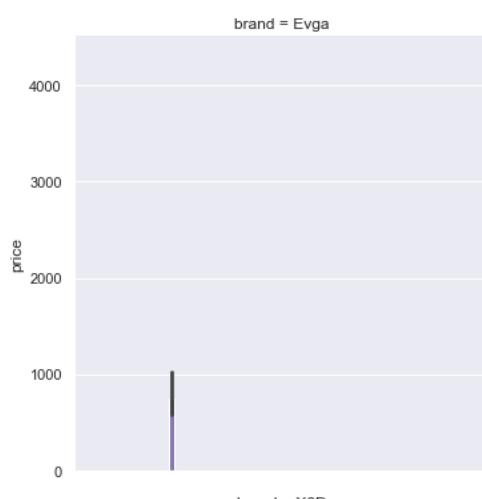
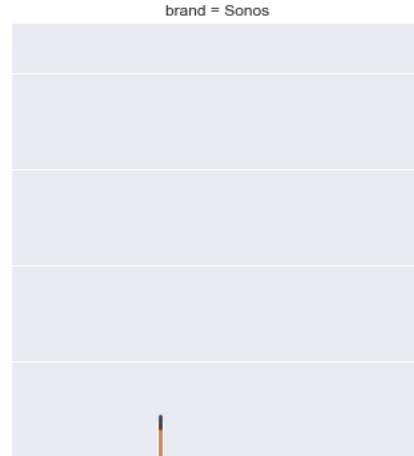
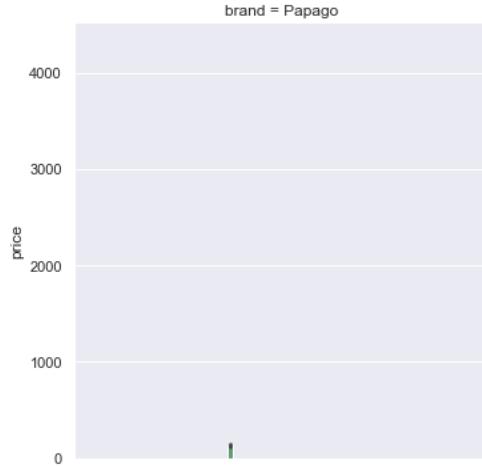
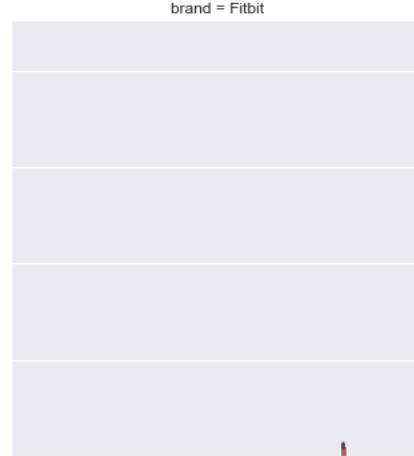
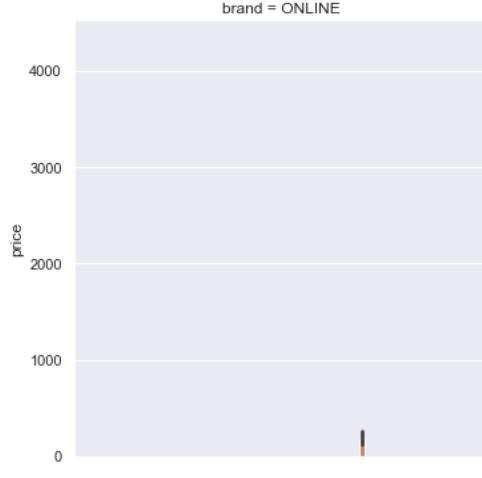
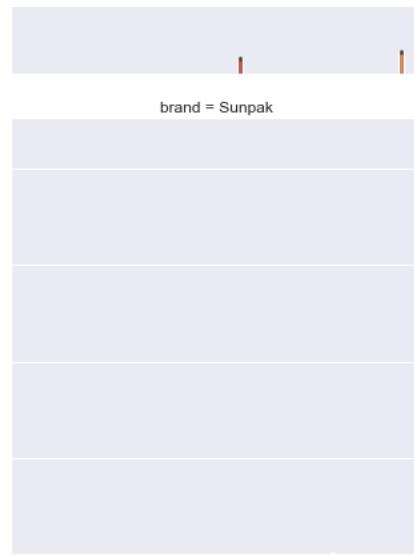
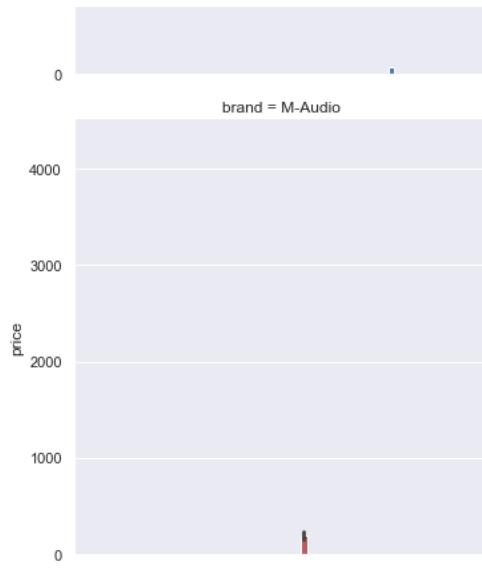


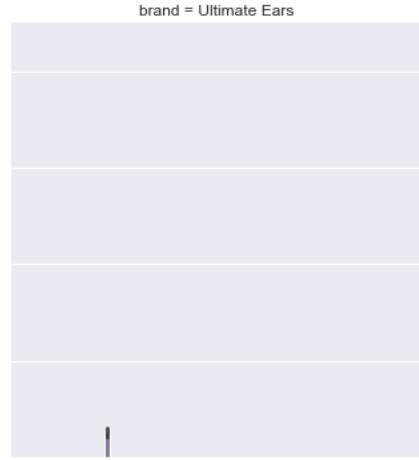
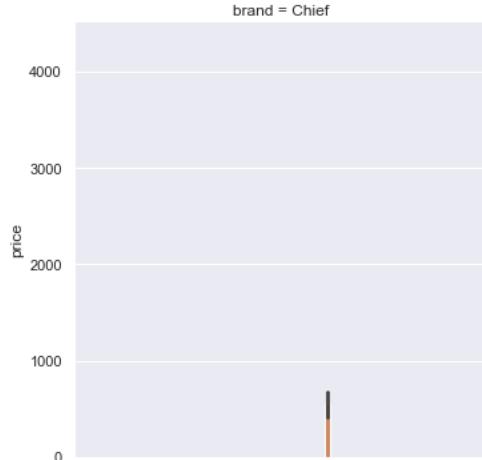
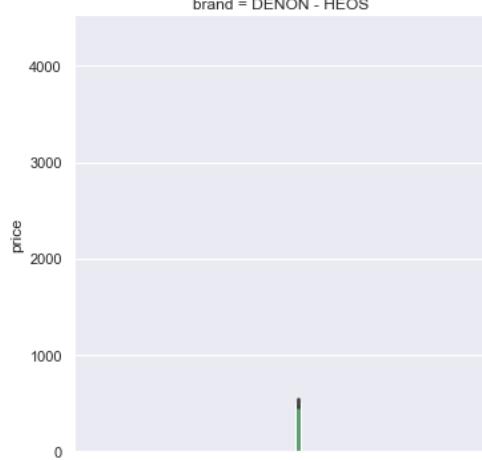
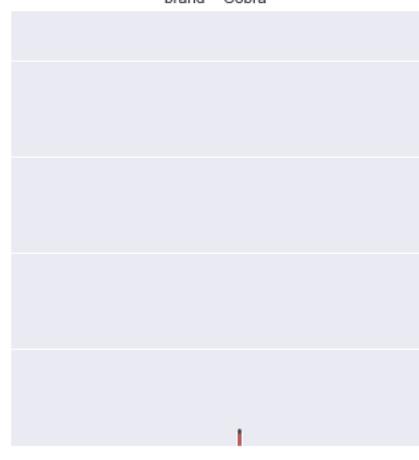
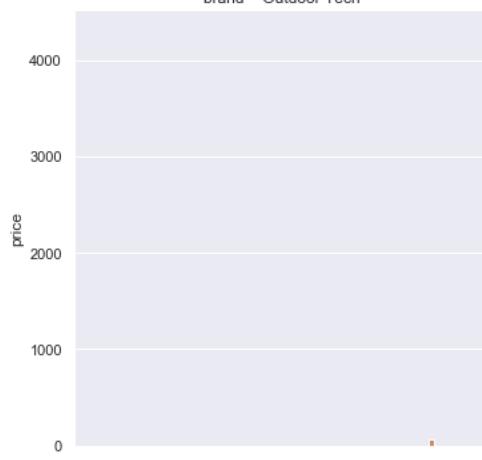


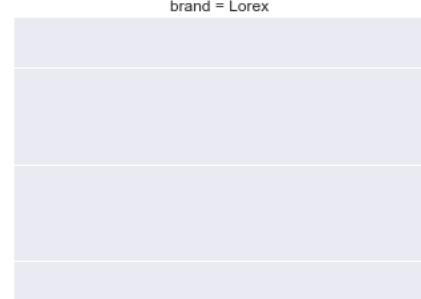
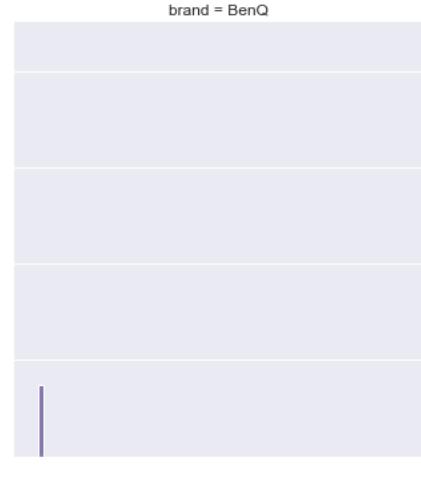
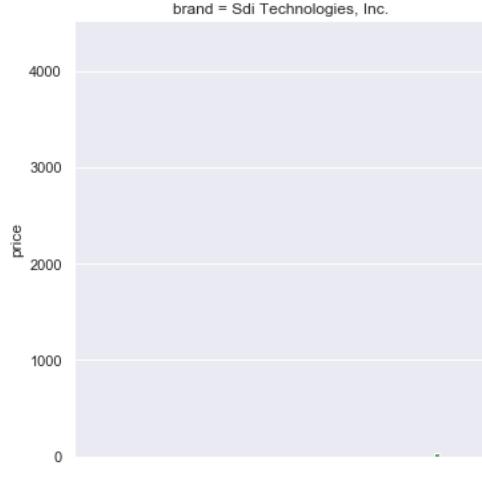
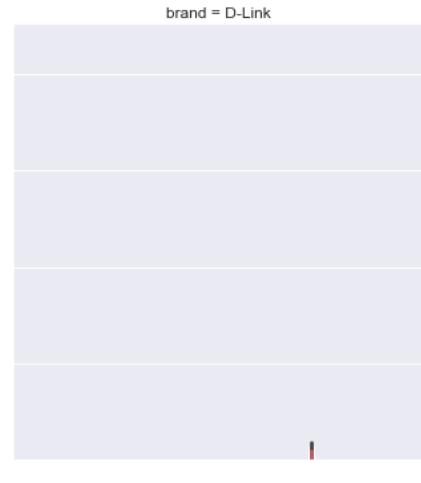
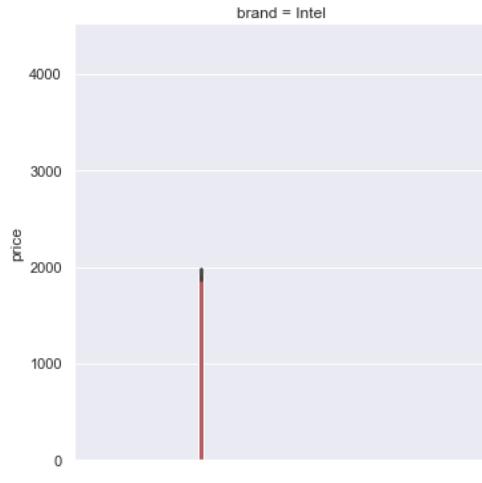
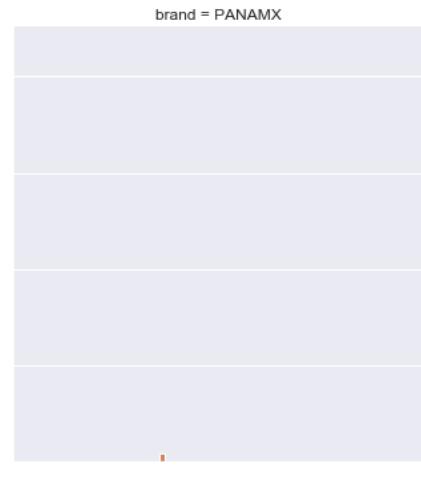
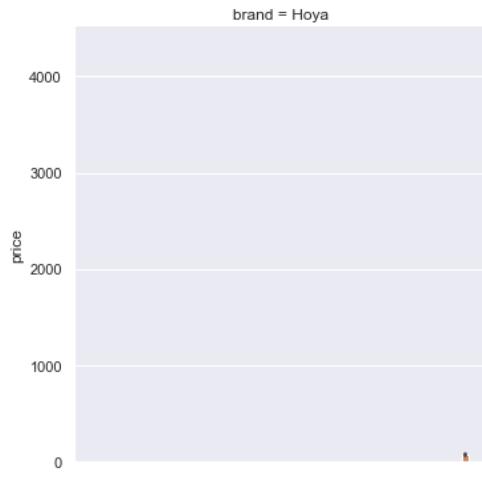


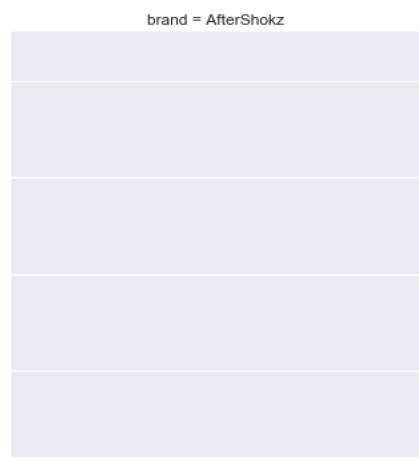
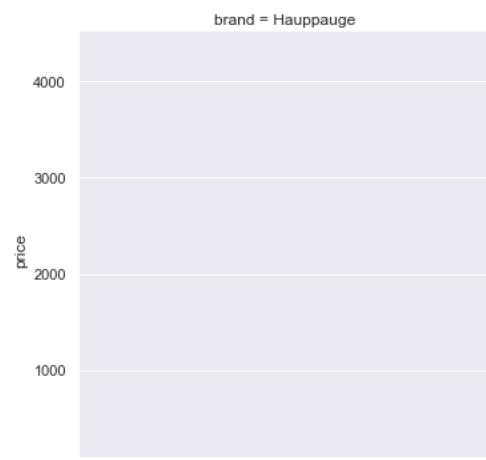
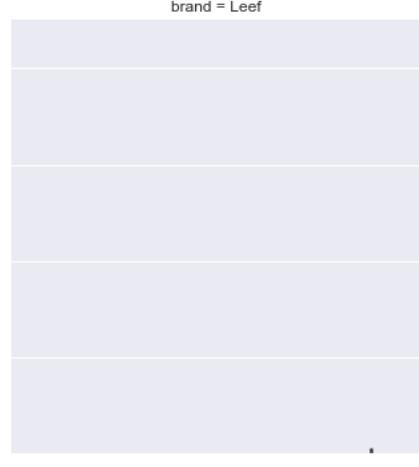
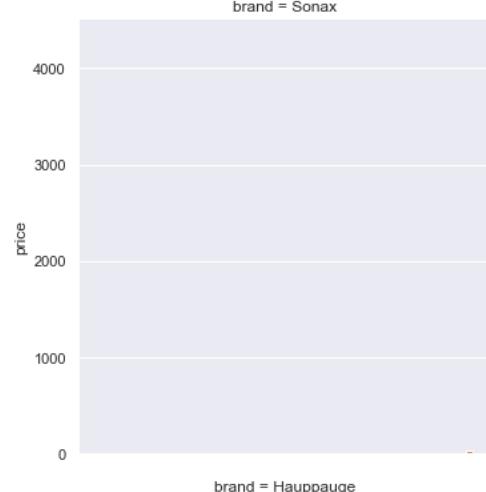
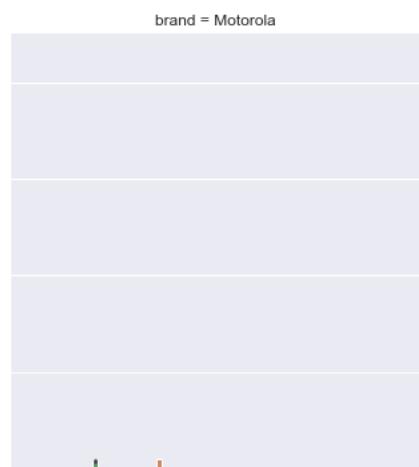
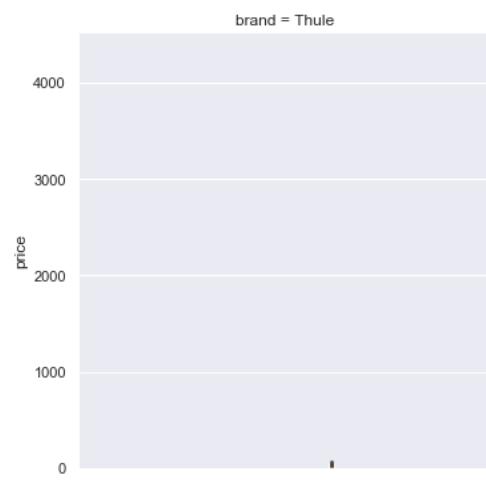
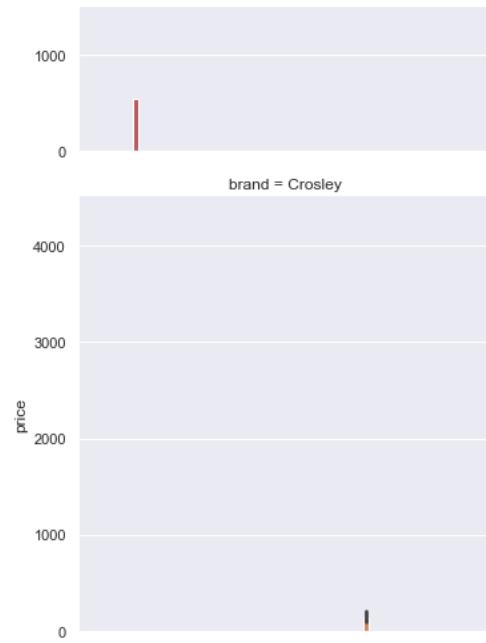


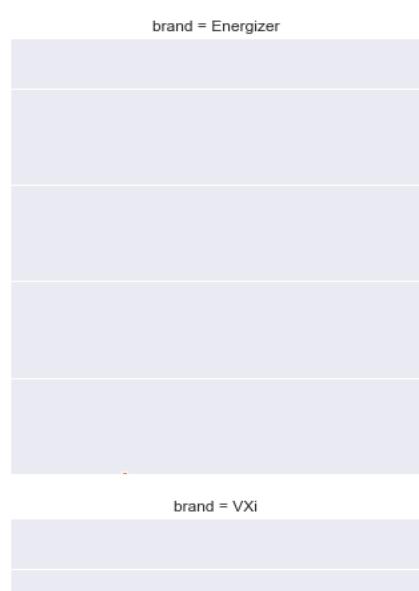
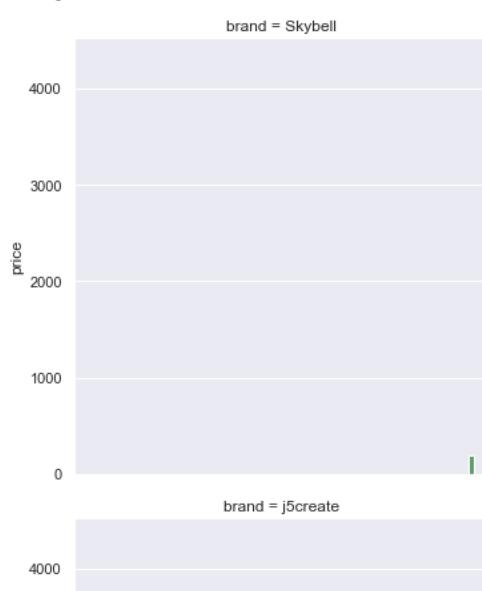
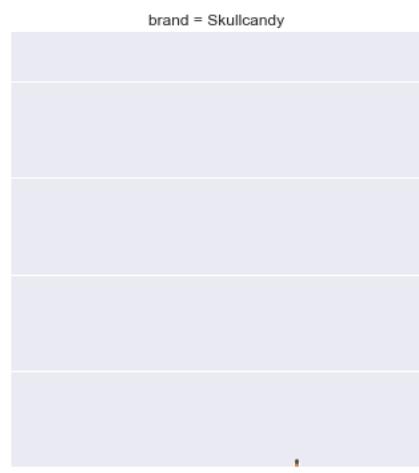
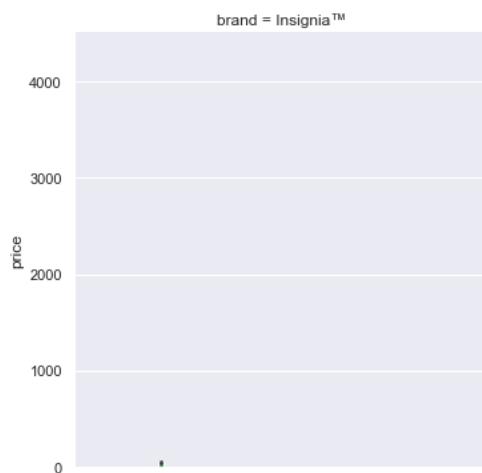
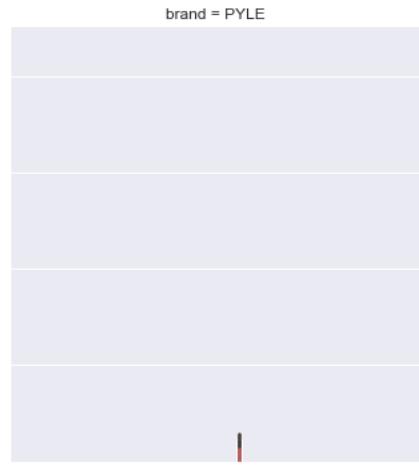
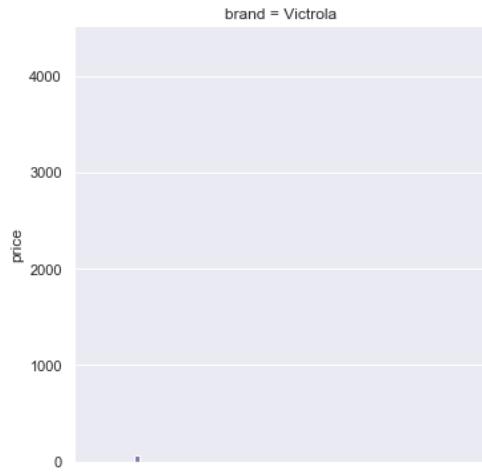
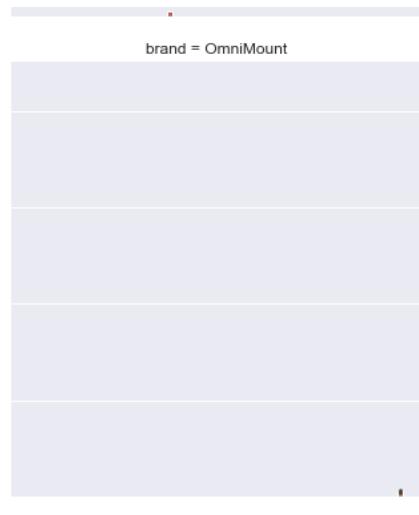
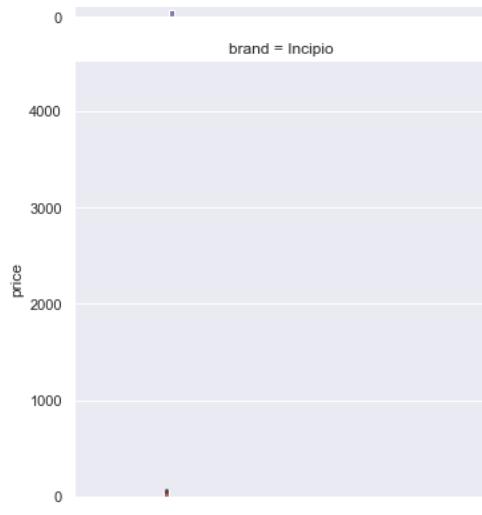






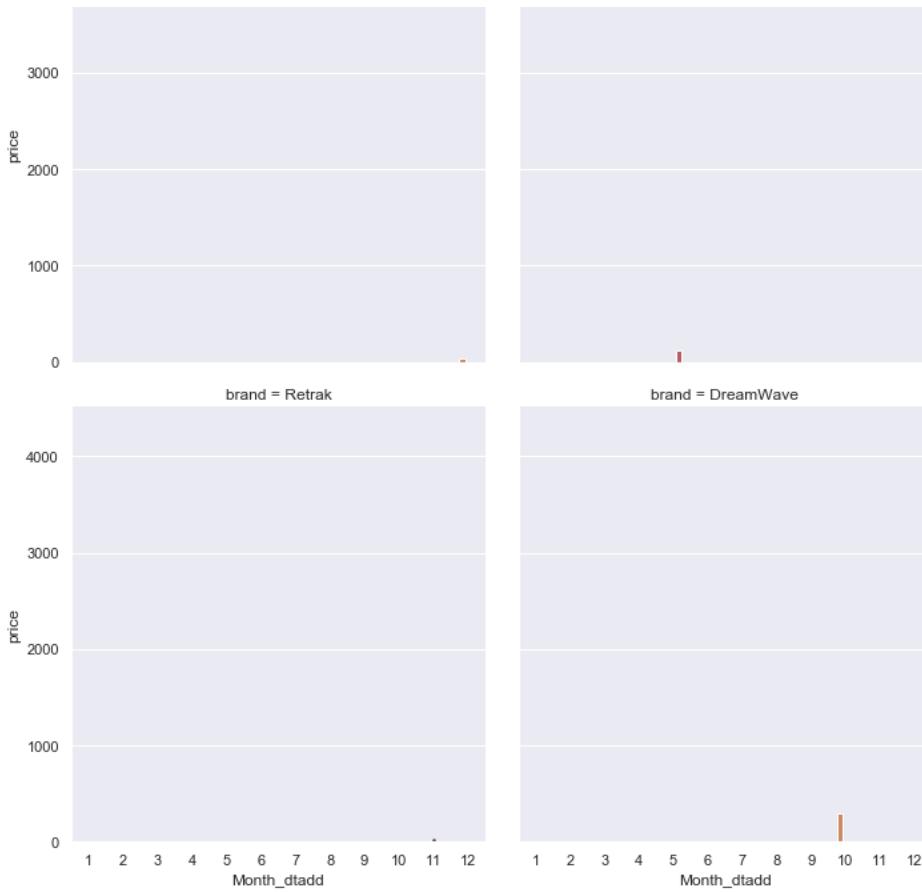






brand = j5create

brand = VXi



Observation

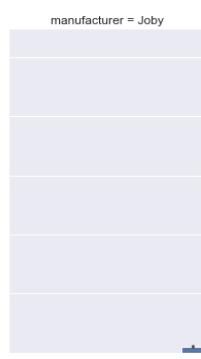
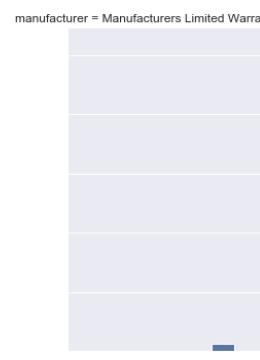
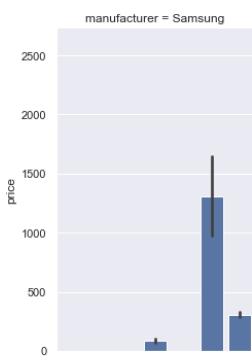
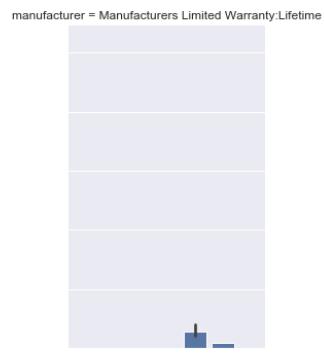
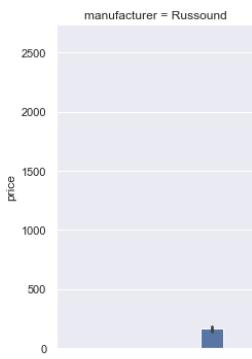
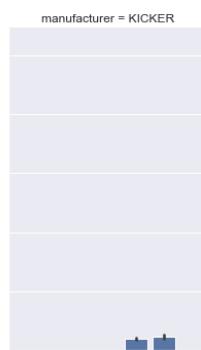
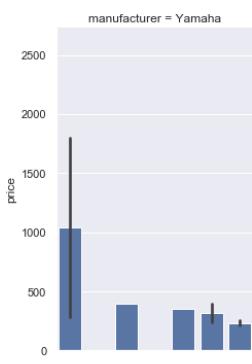
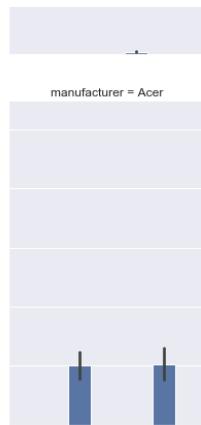
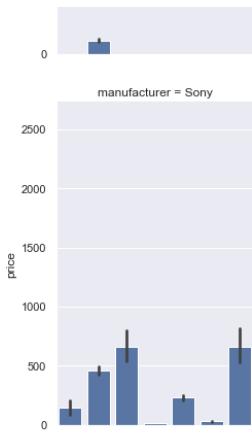
From this graph I can conclude Sony has more price than other brand
I can also conclude that in 4th month of 2017 price is highest

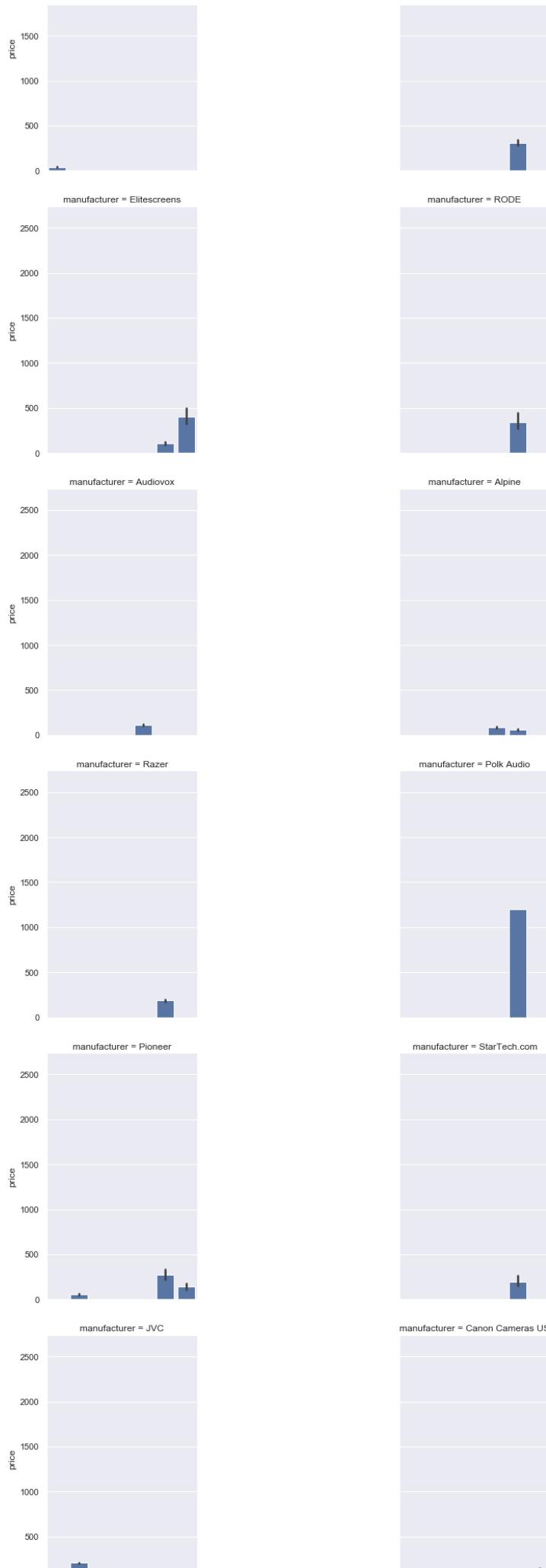
I will plot manufacturer with dateUpdated

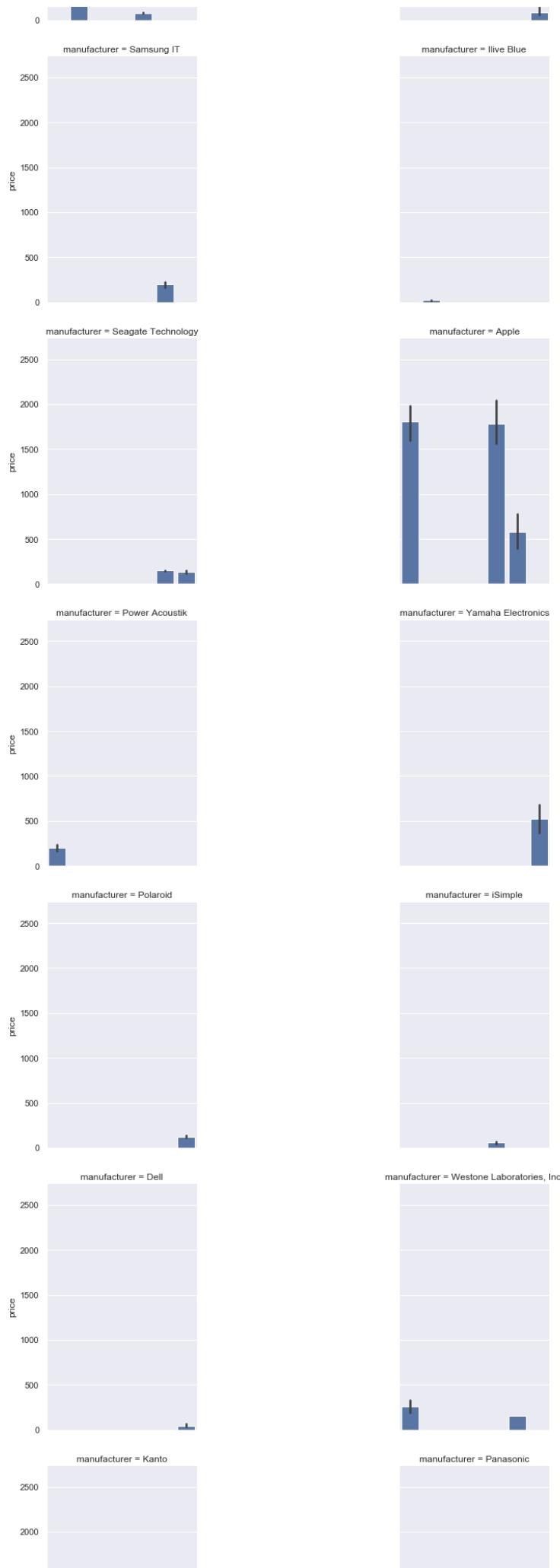
In [75]:

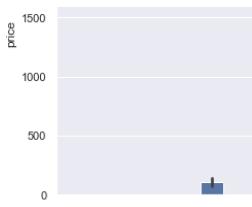
```
sns.set(style='darkgrid', font_scale=1.0)
sns.catplot(x='Month_updated',
            y='price', hue='Year_updated',
            col='manufacturer',
            col_wrap=2,
            data=dt_electronics,
            kind='bar')
plt.savefig('Dateupdated_manufacturer.png')
```



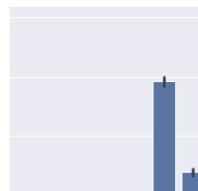




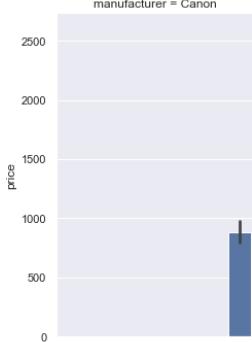




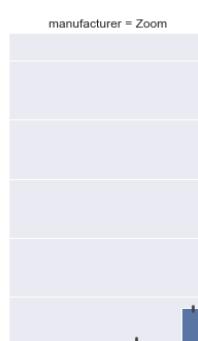
manufacturer = Canon



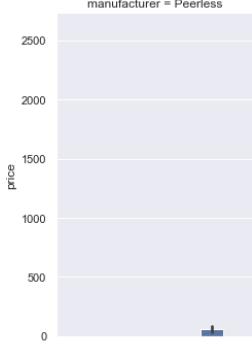
manufacturer = Corsair



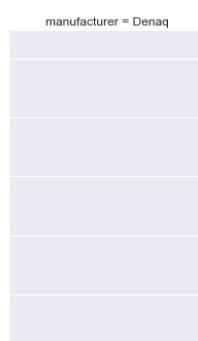
manufacturer = Peerless



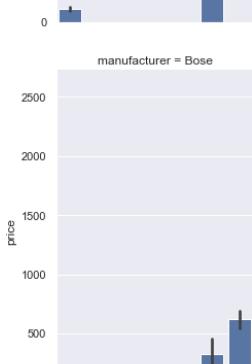
manufacturer = Zoom



manufacturer = Sennheiser



manufacturer = Denaq



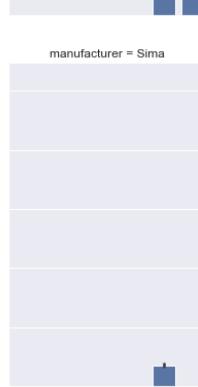
manufacturer = Bose



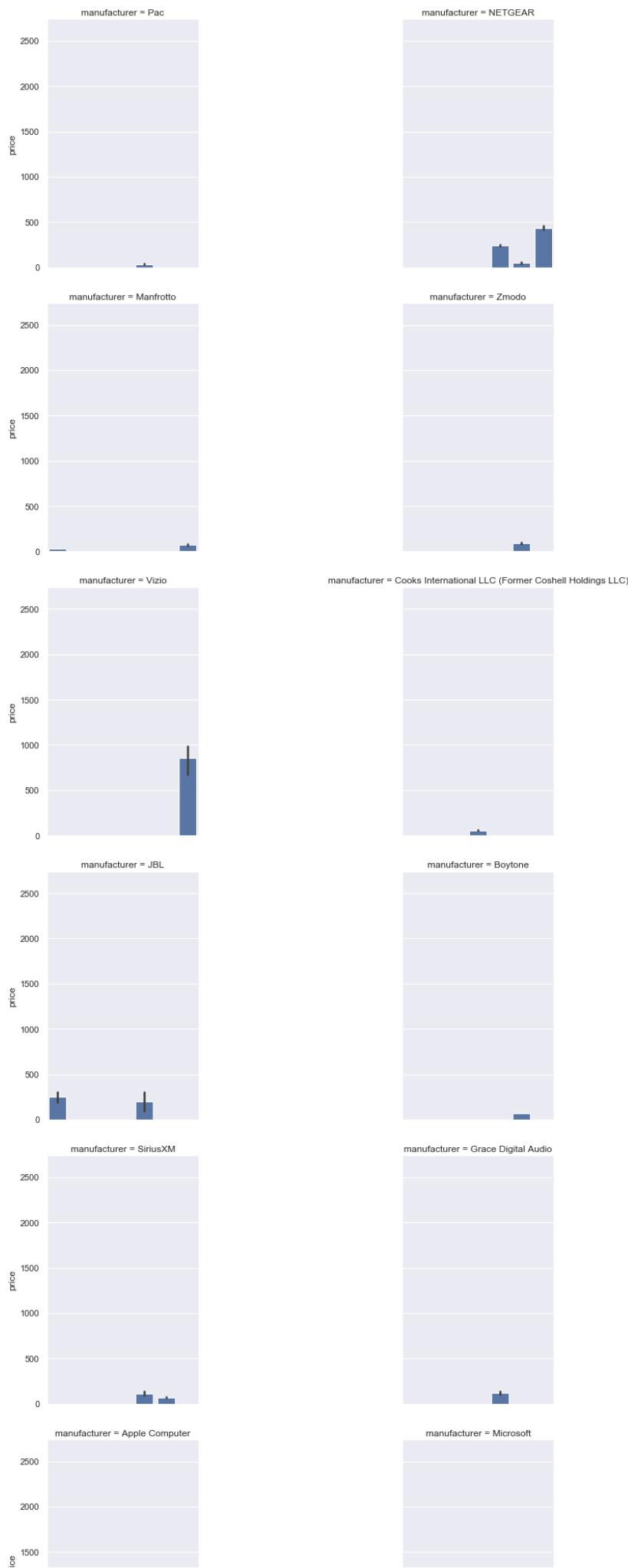
manufacturer = Western Digital

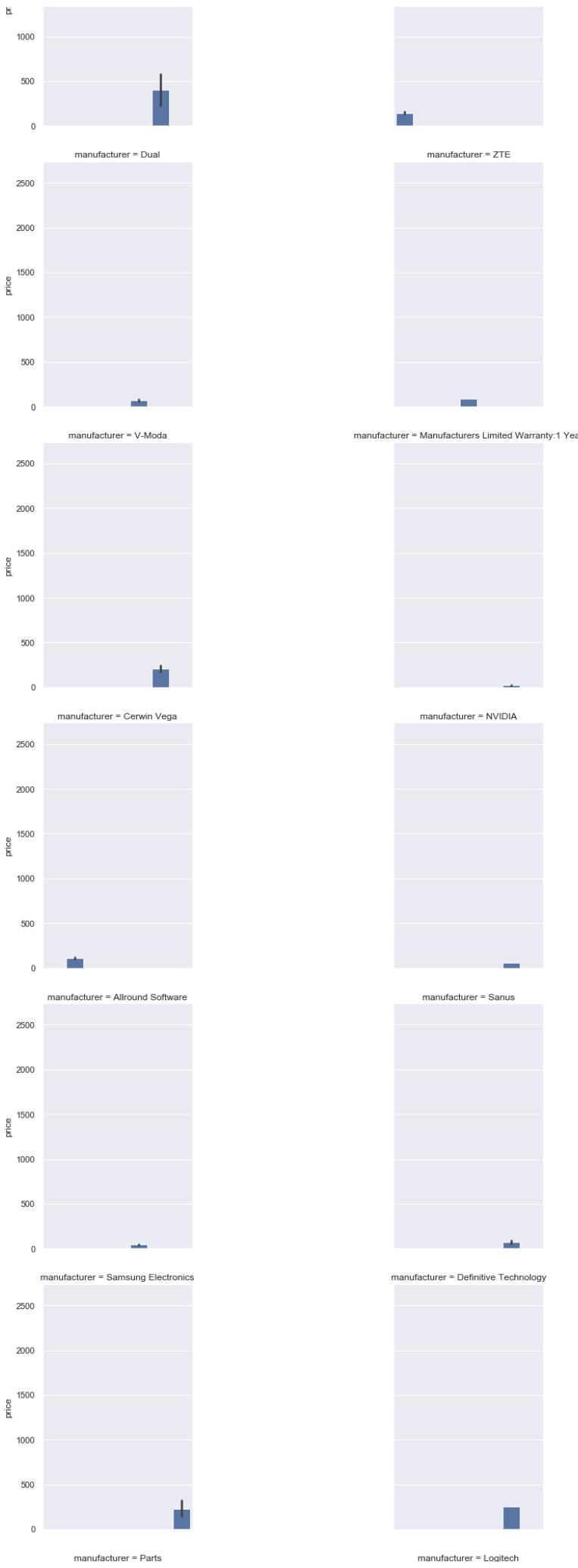


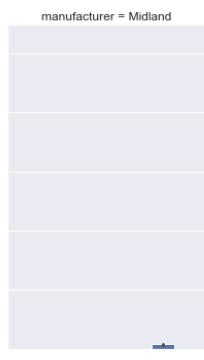
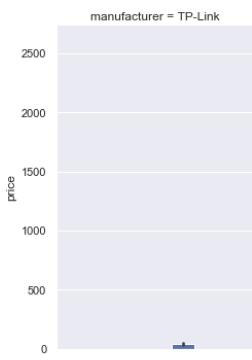
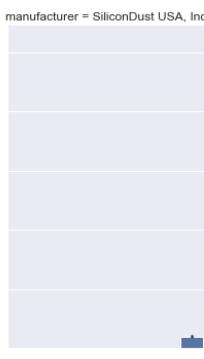
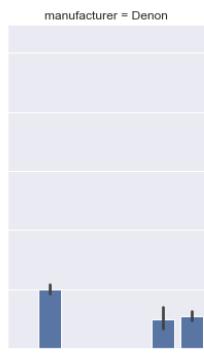
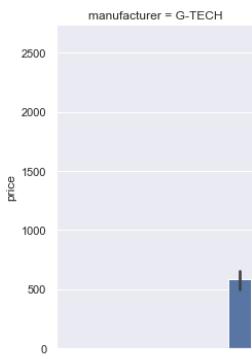
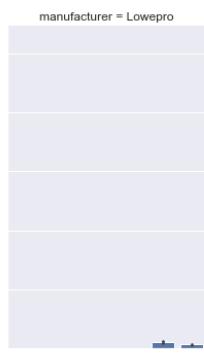
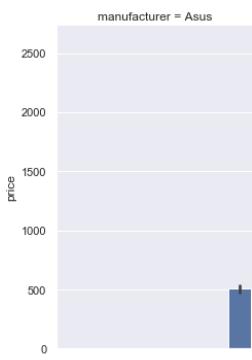
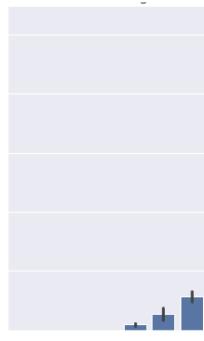
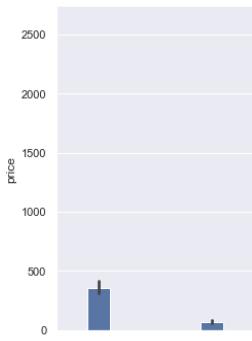
manufacturer = Slingbox

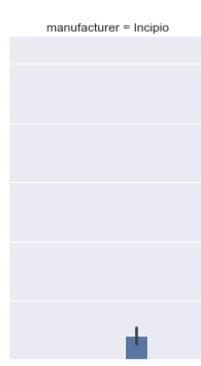
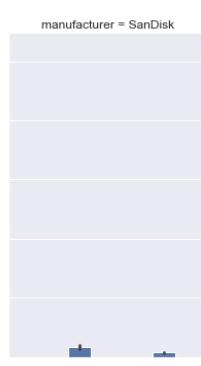
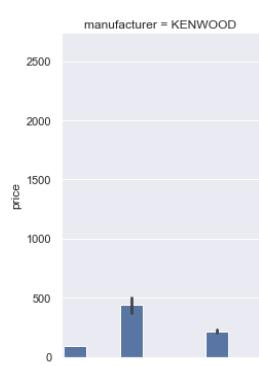
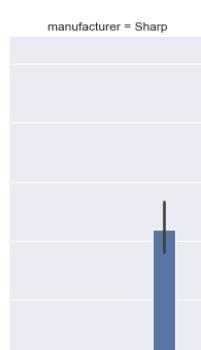
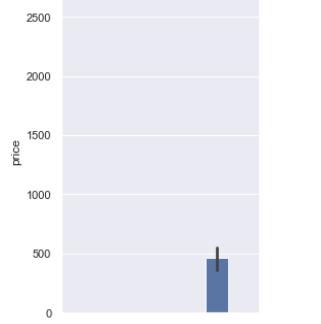


manufacturer = Sima



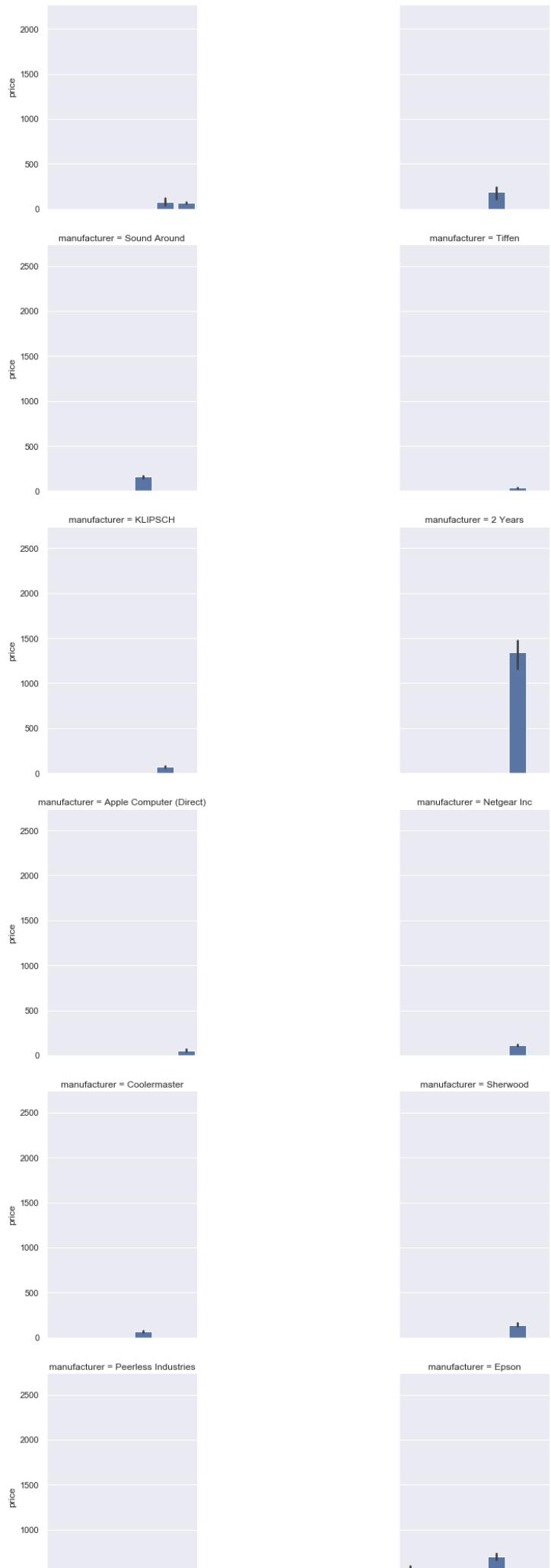


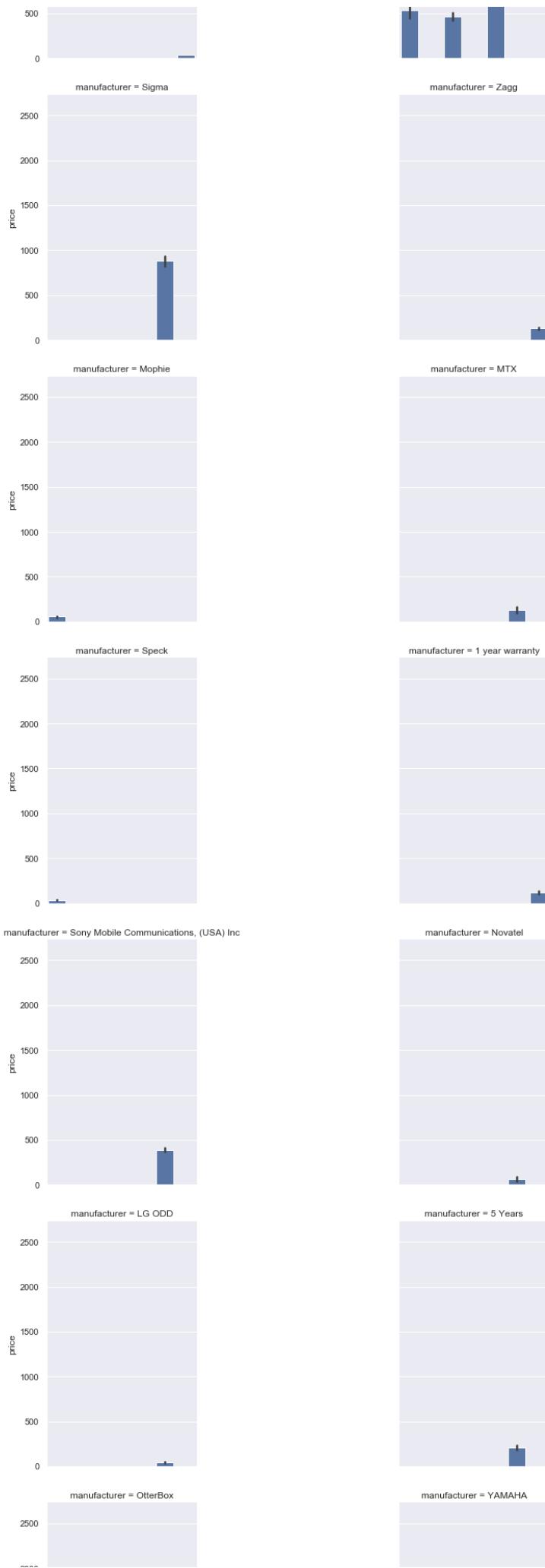


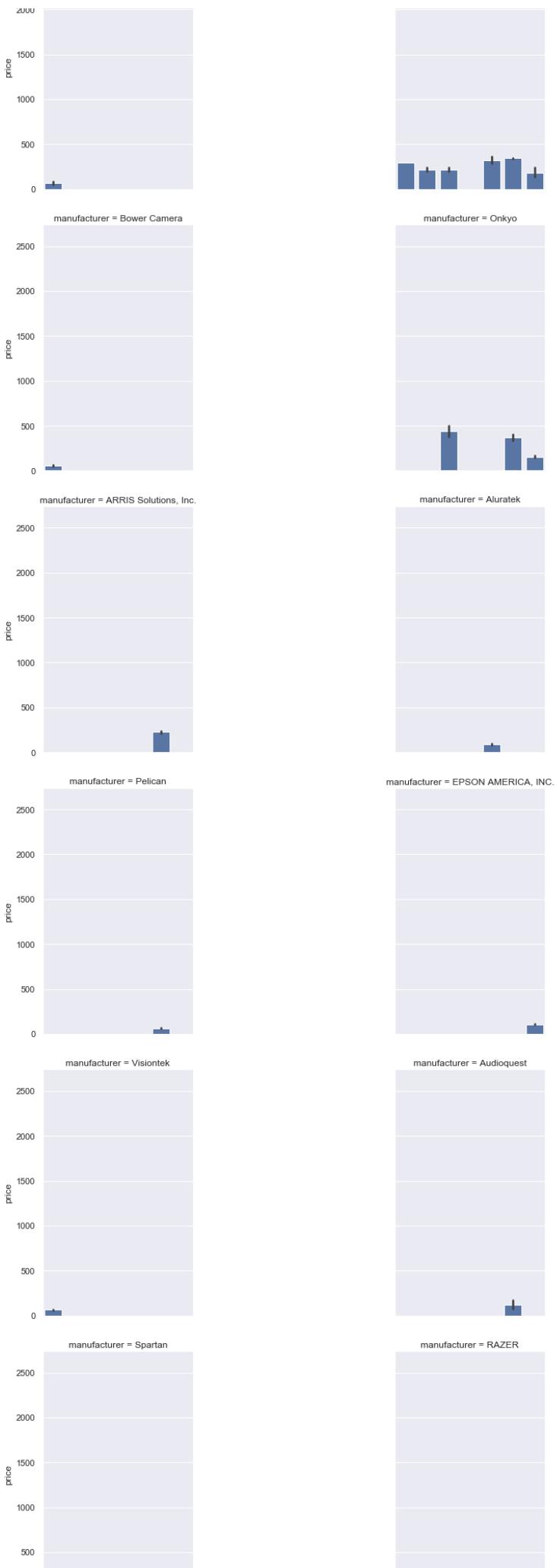


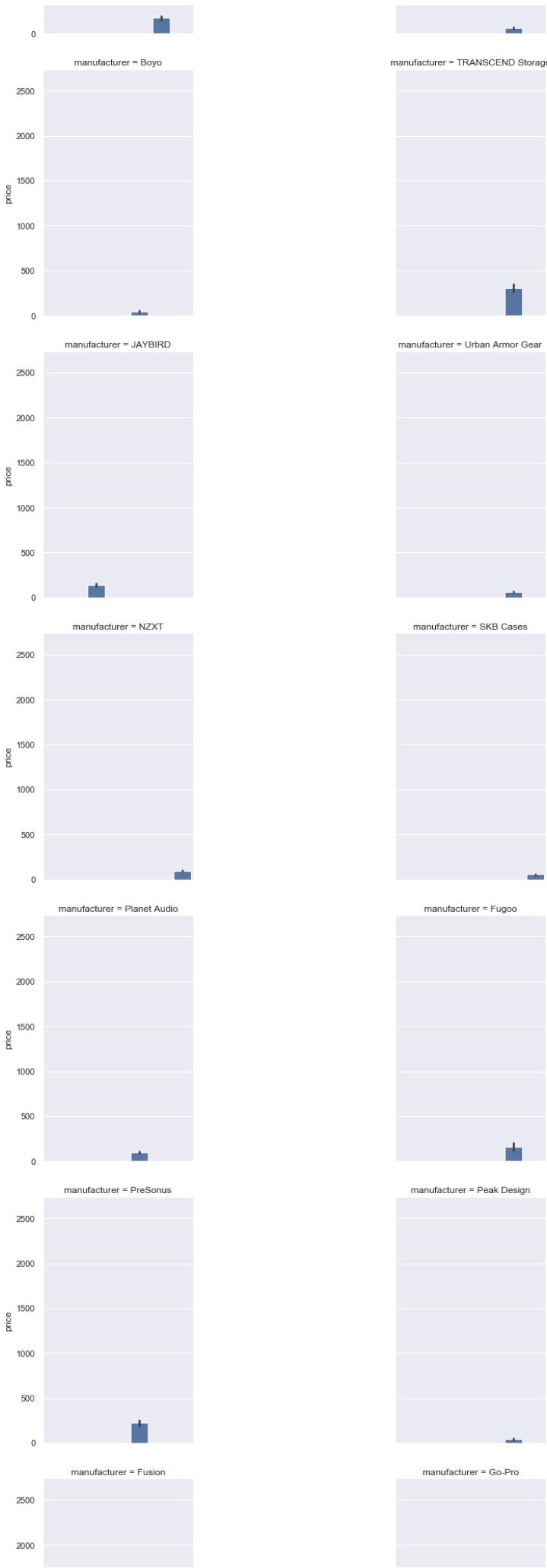


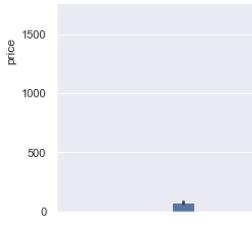




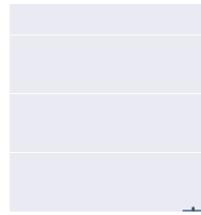




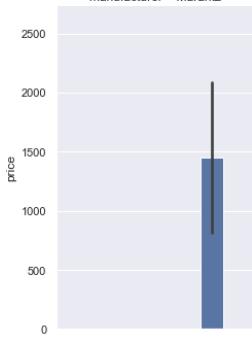




manufacturer = Marantz

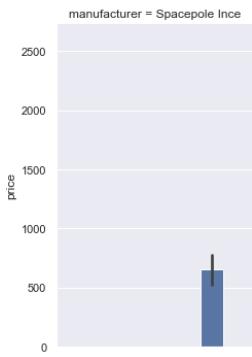


manufacturer = Sunpak



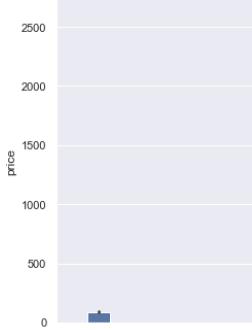
manufacturer = Spacepole Inc

manufacturer = Siriusxm



manufacturer = Atrend

manufacturer = HP Inc



manufacturer = MEE audio

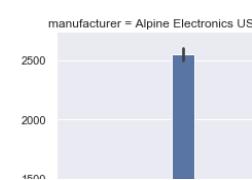
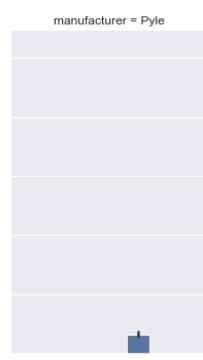
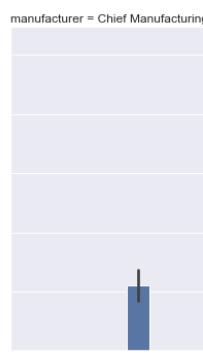
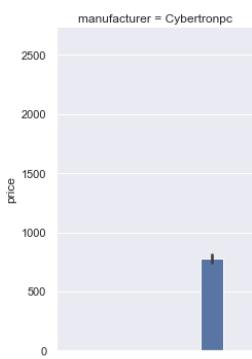
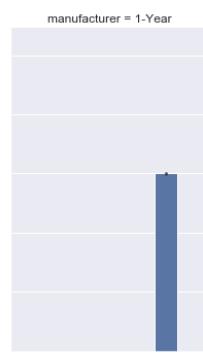
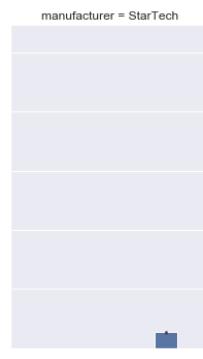
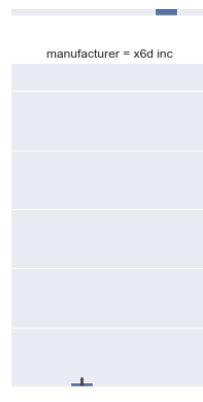
manufacturer = Papago

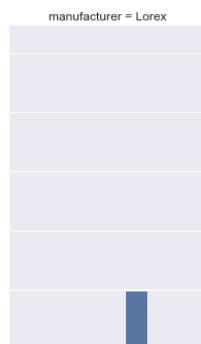
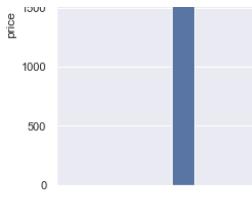


manufacturer = Sonos

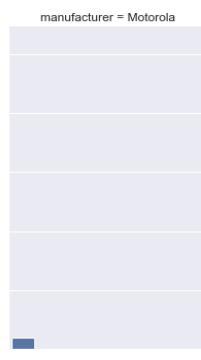
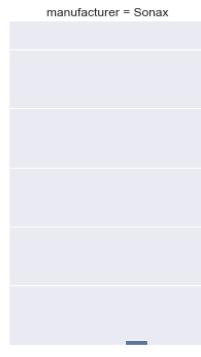
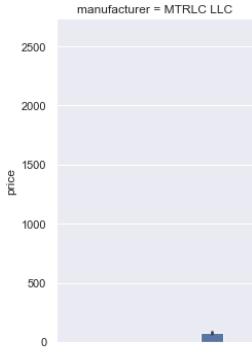
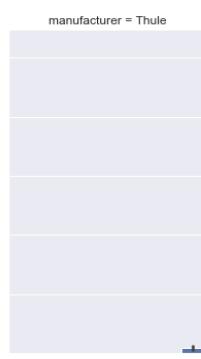
manufacturer = ELITE SCREENS DIRECTSHIP

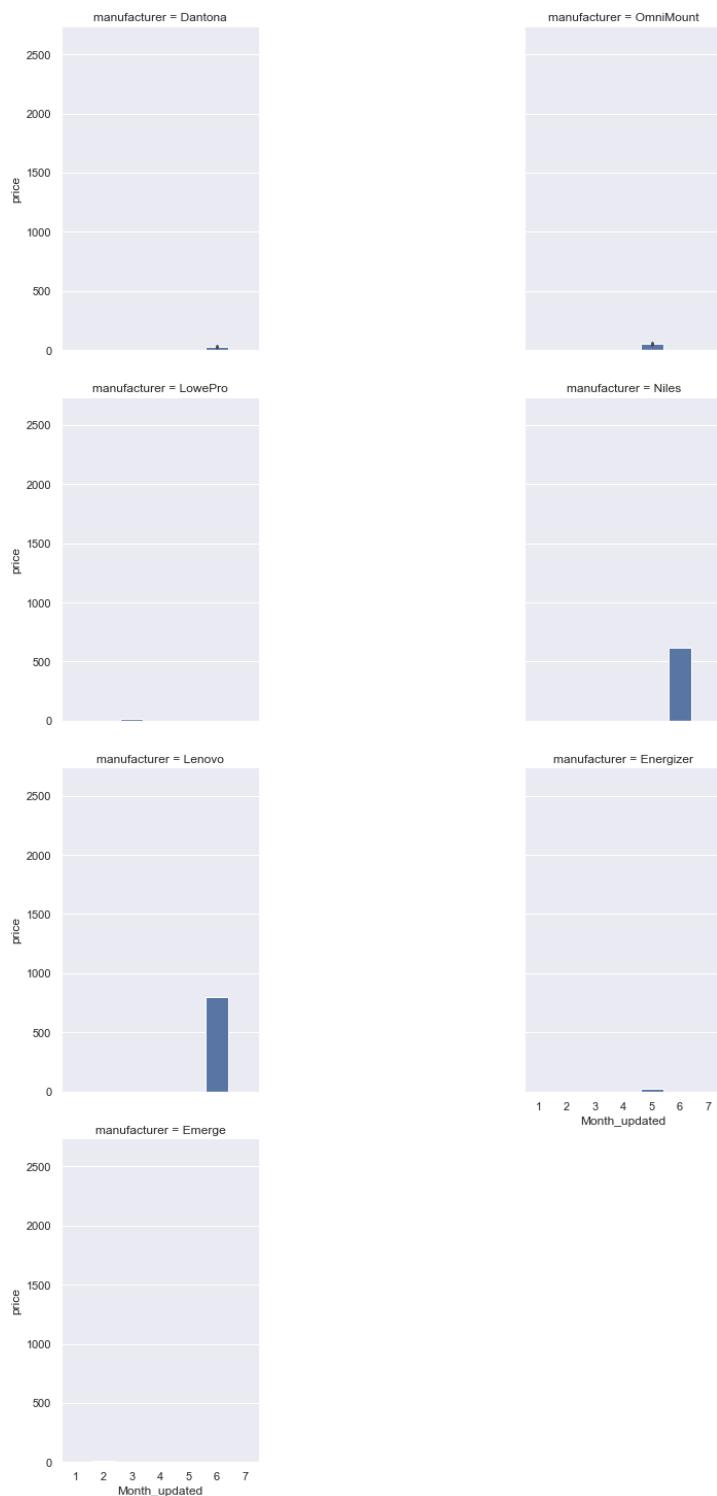






manufacturer = Jvc Real Sound System Z Series Ha-Sz2000 Stream Woofer Db Headphones (black, Japan Import) V-Moda





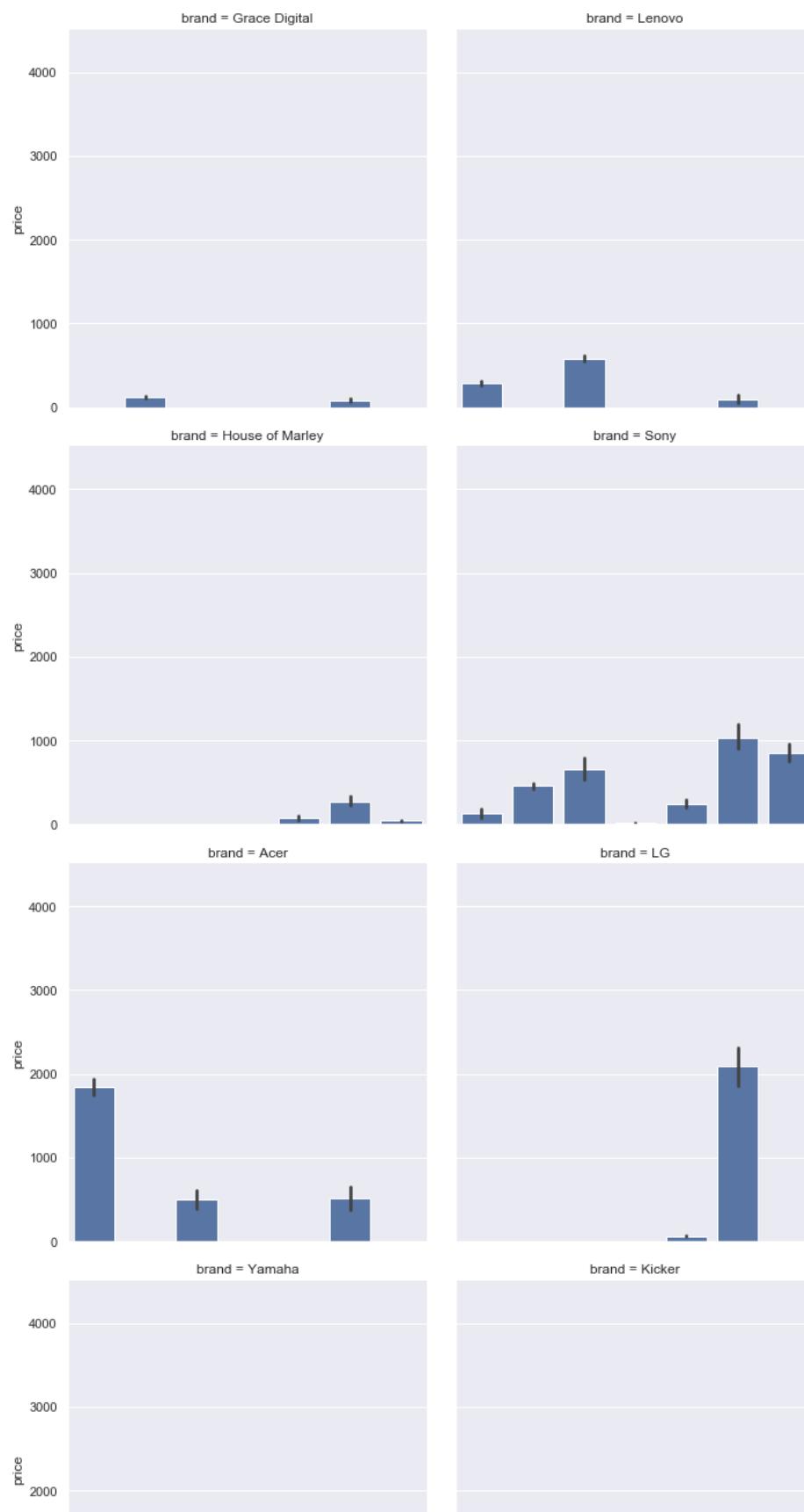
Observation

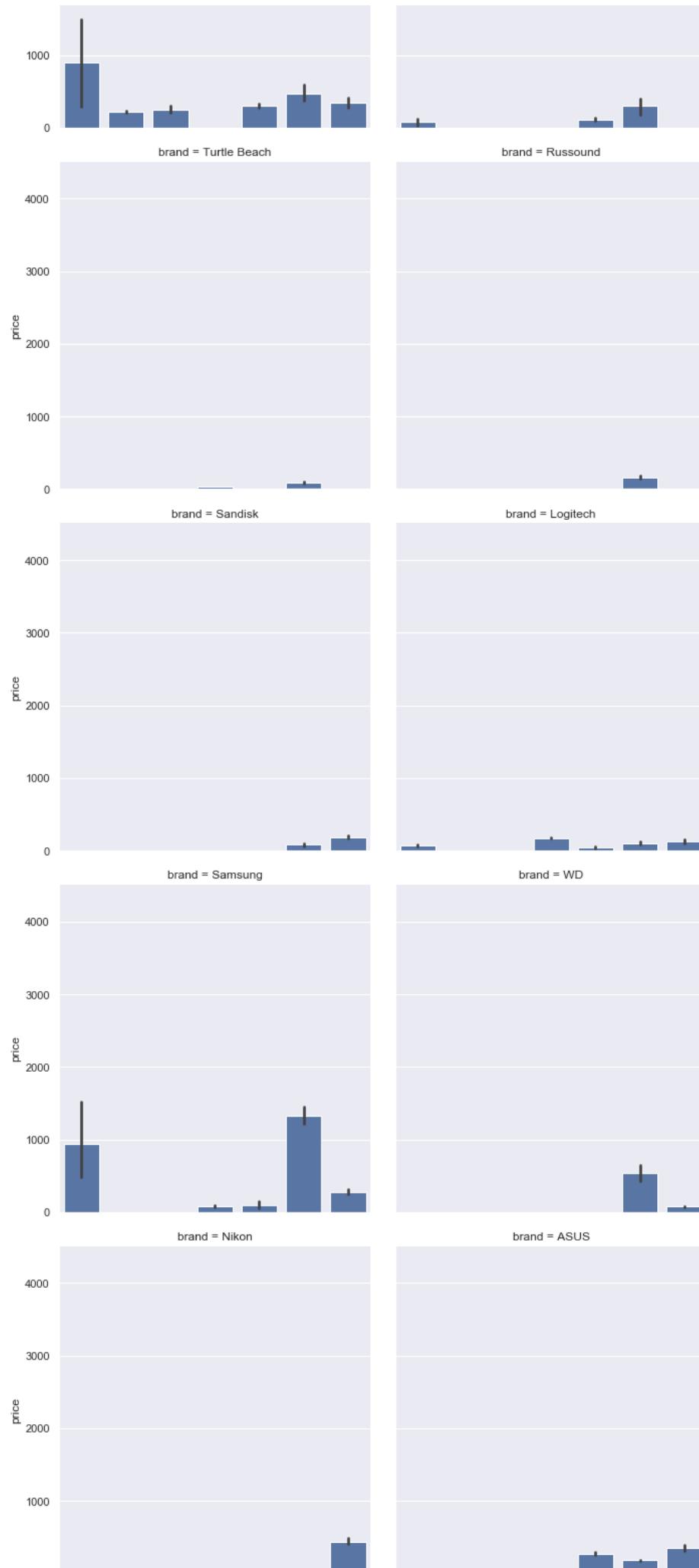
From this graph I can conclude Apple has more price than other category
 I can also conclude that in 1st and 5th month of 2018 price is highest

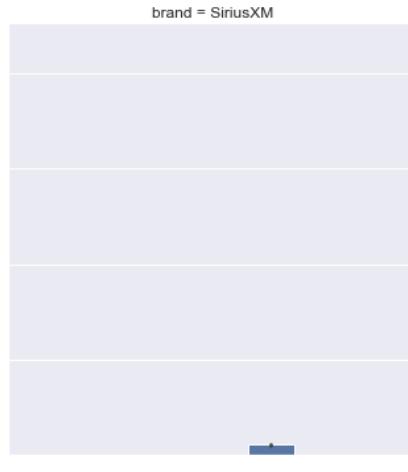
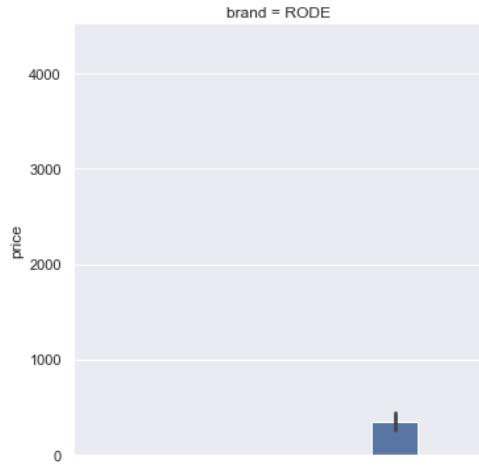
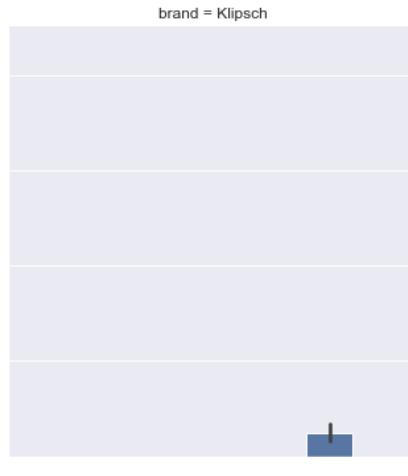
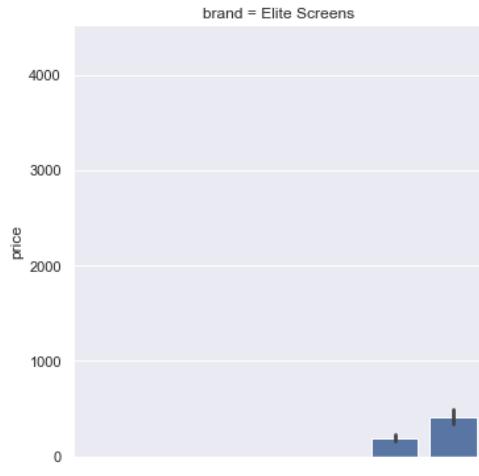
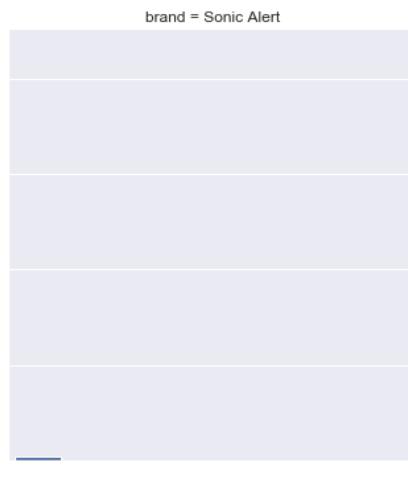
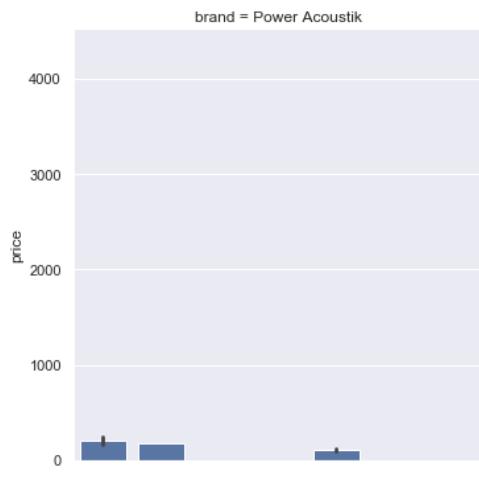
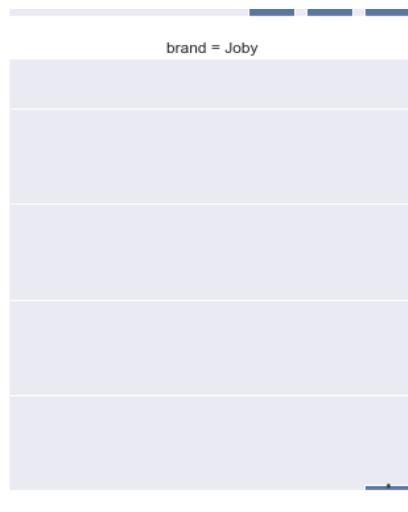
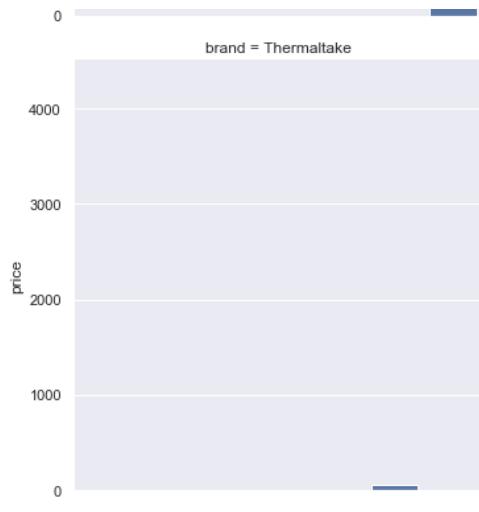
I will plot band with dateupdated

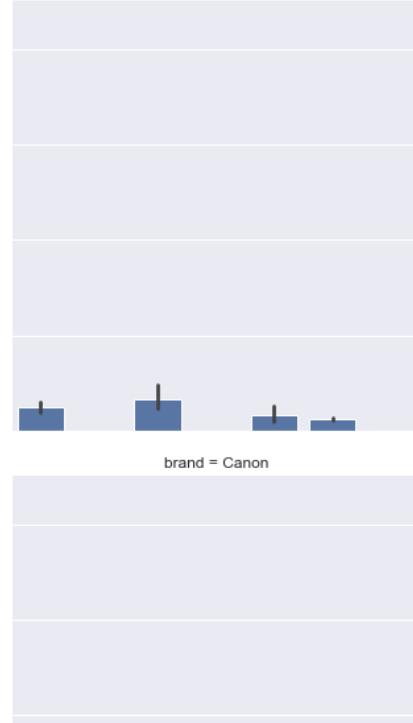
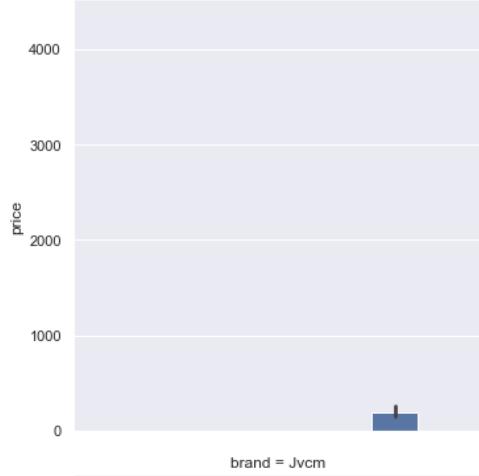
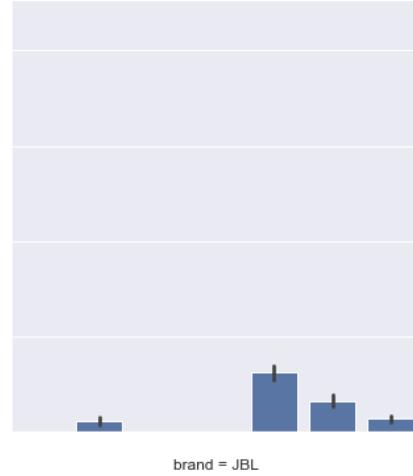
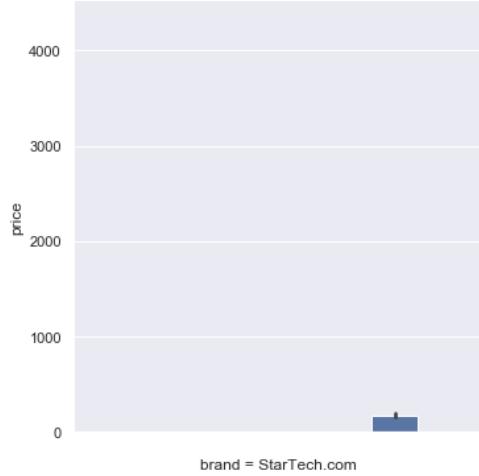
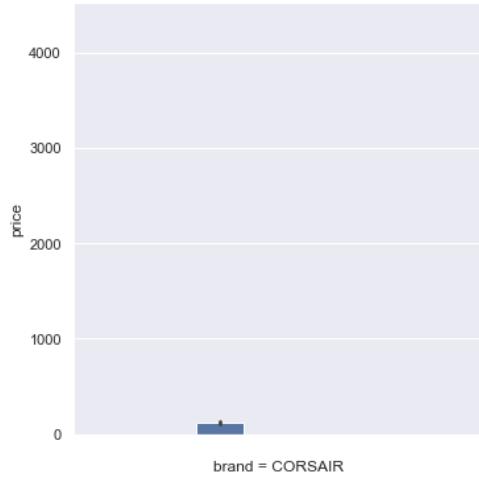
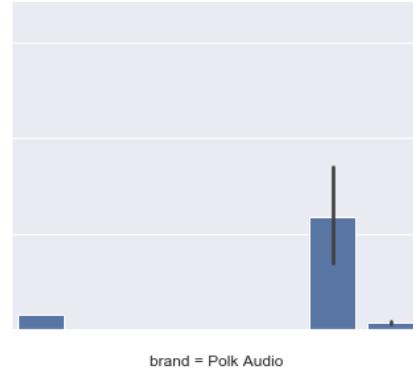
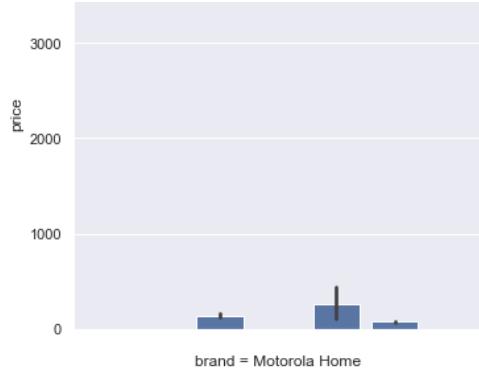
In [76]:

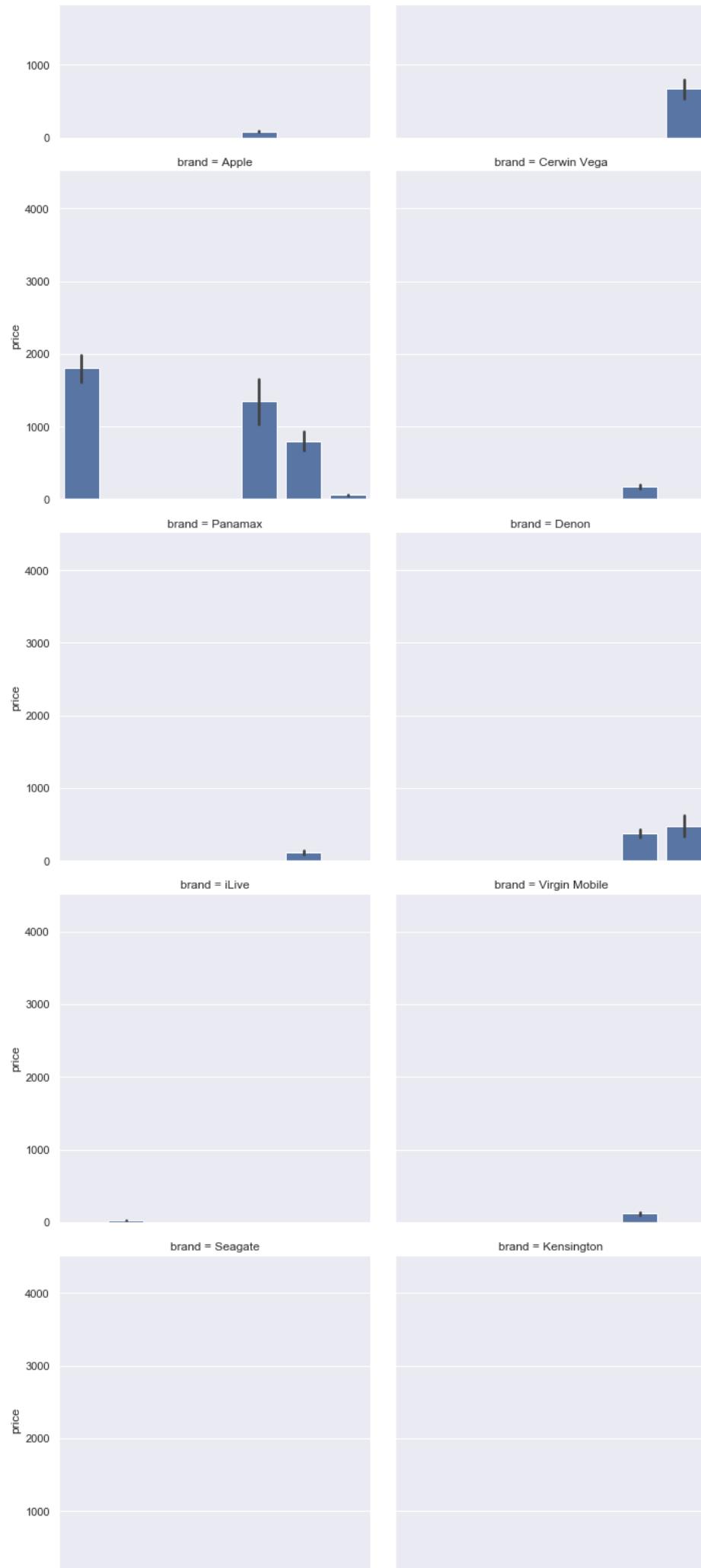
```
sns.set(style='darkgrid', font_scale=1.0)
sns.catplot(x='Month_updated',
            y='price', hue='Year_updated',
            col='brand',
            col_wrap=2,
            data=dt_electronics,
            kind='bar')
plt.savefig('Dateupdated_brand.png')
```

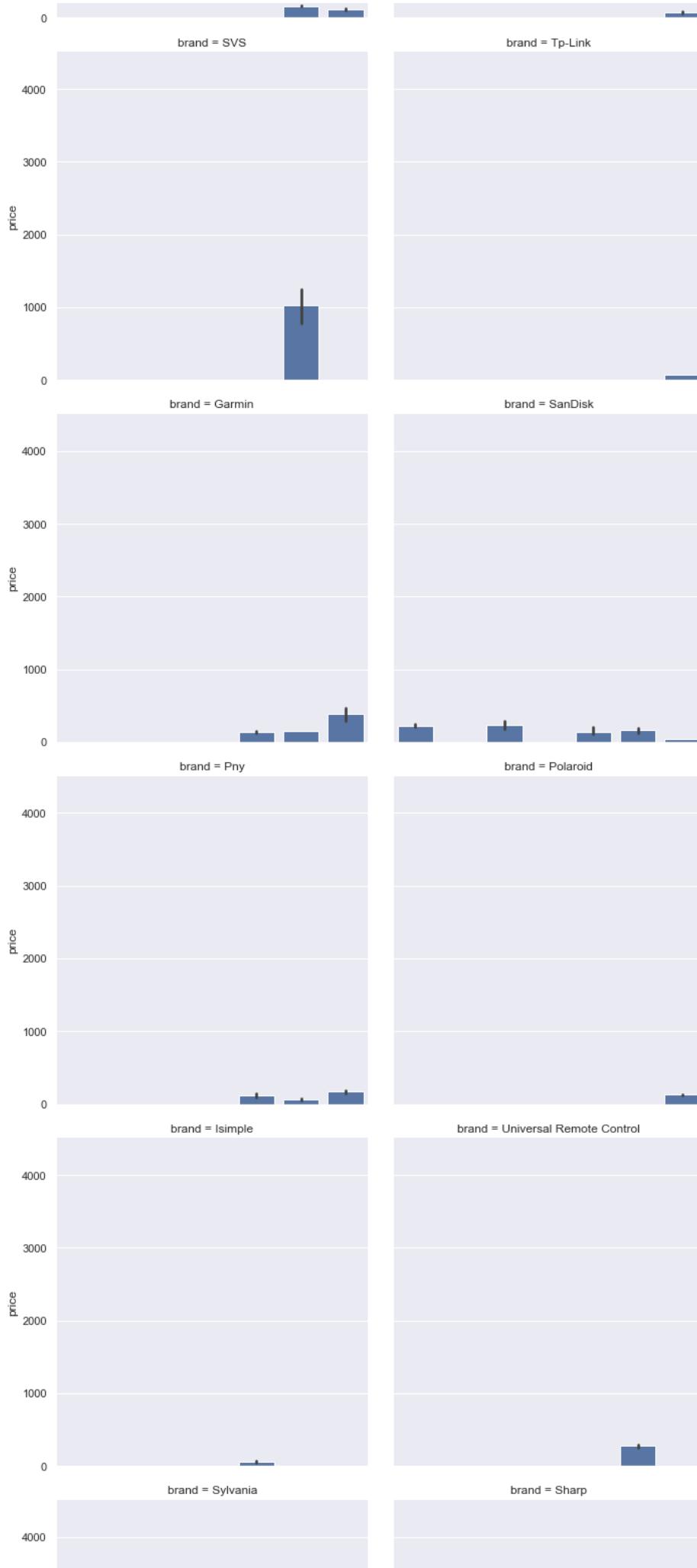


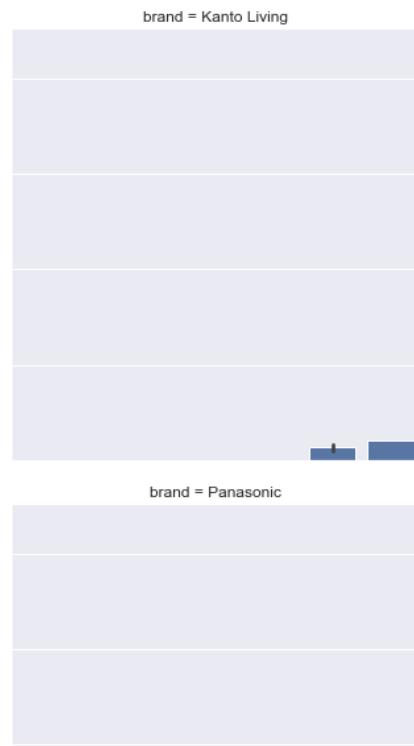
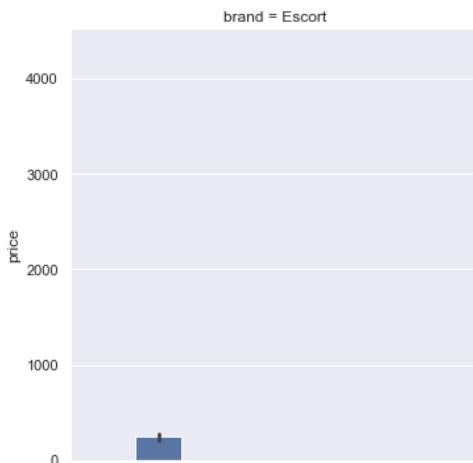
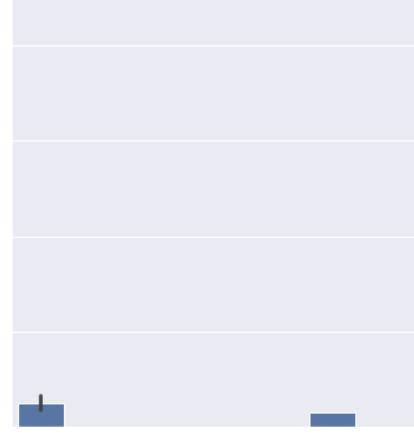
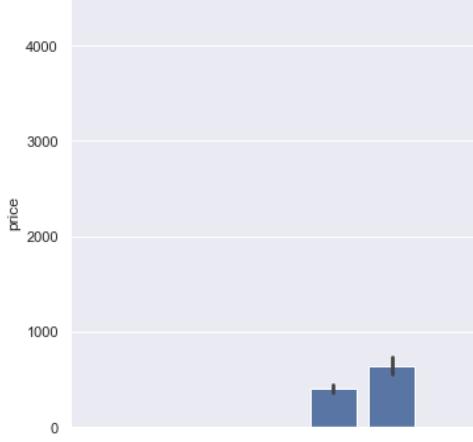
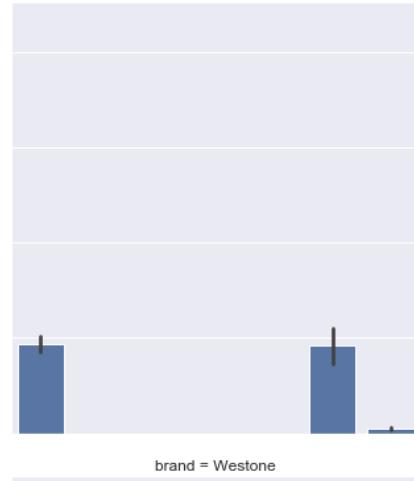
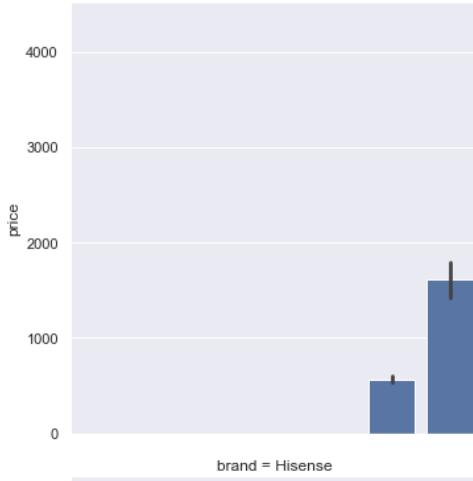
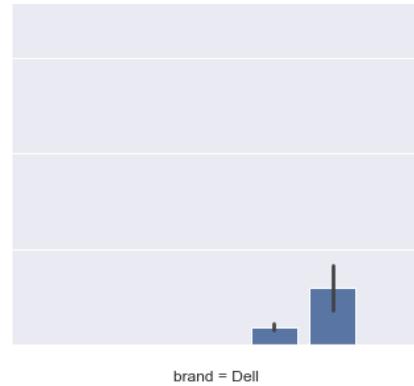
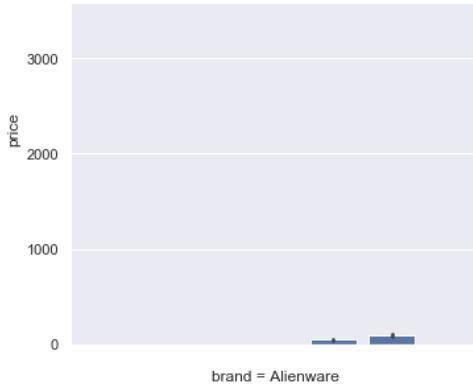




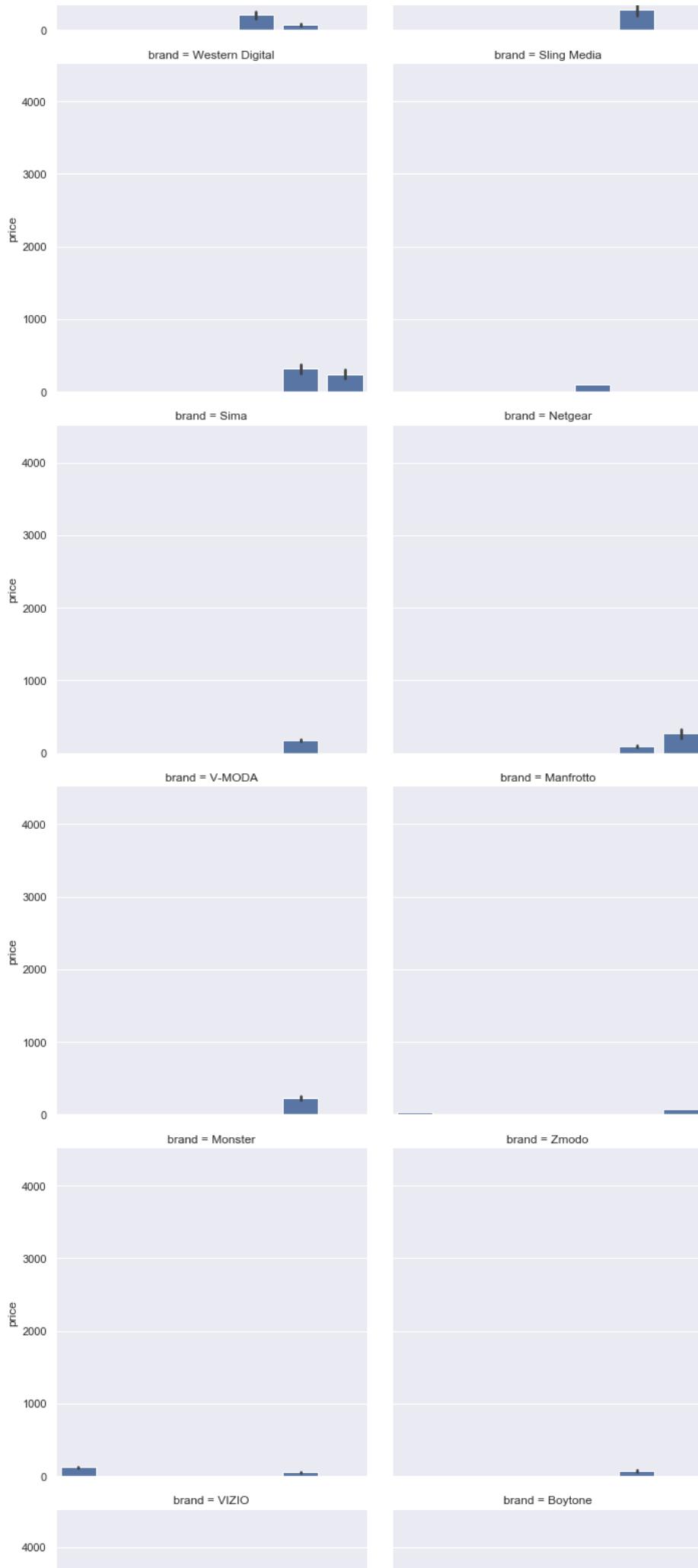


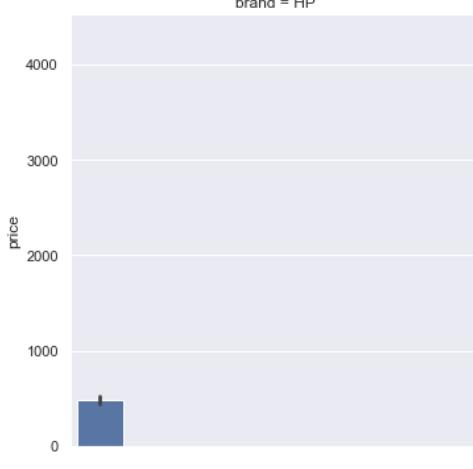
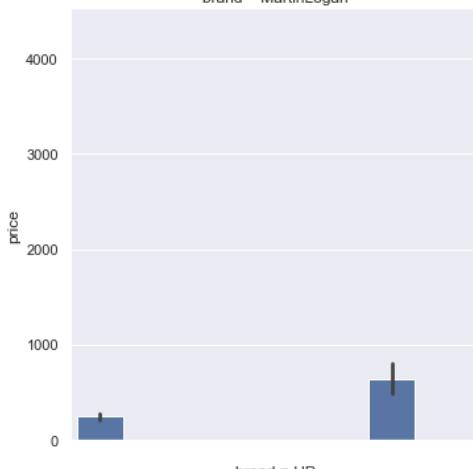
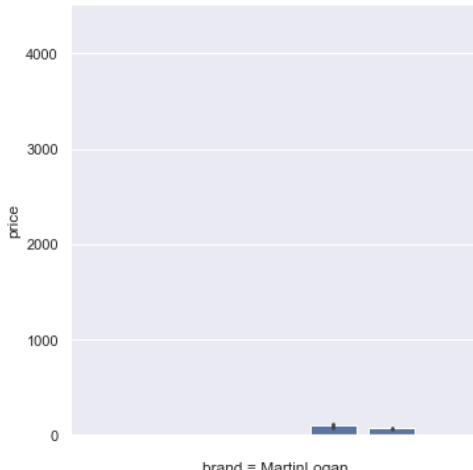
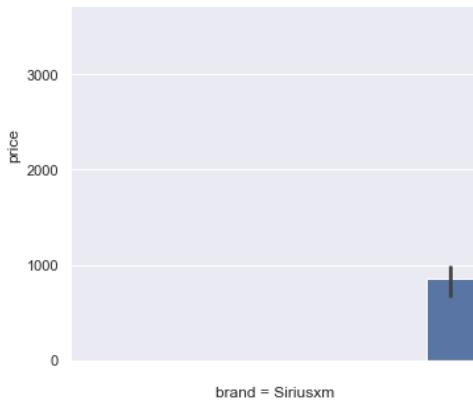


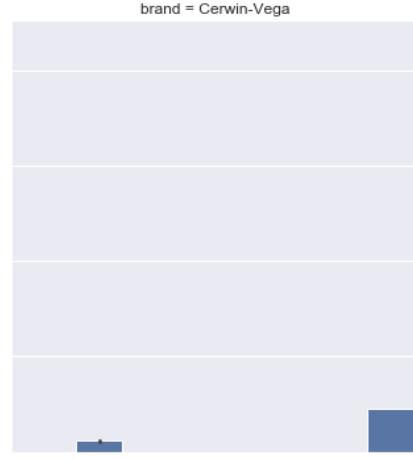
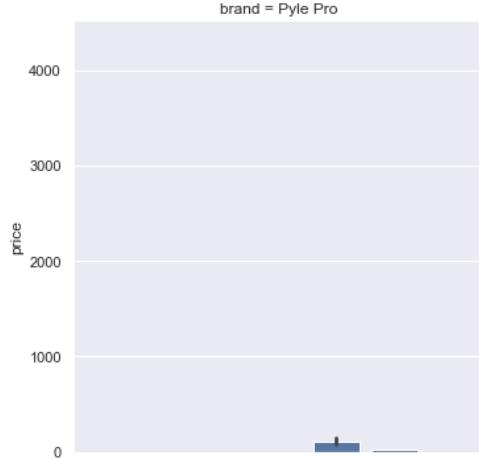
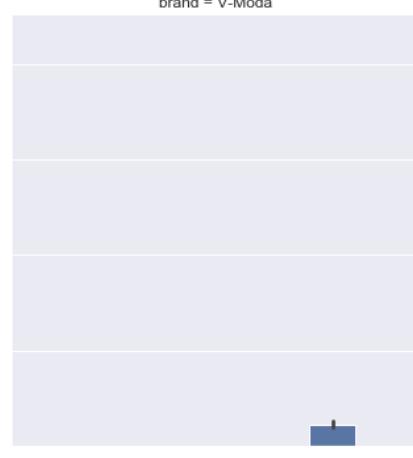
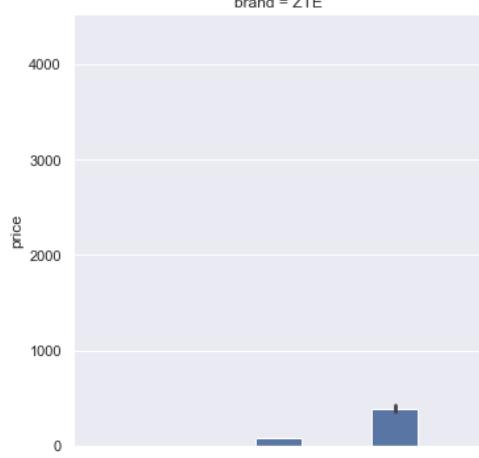
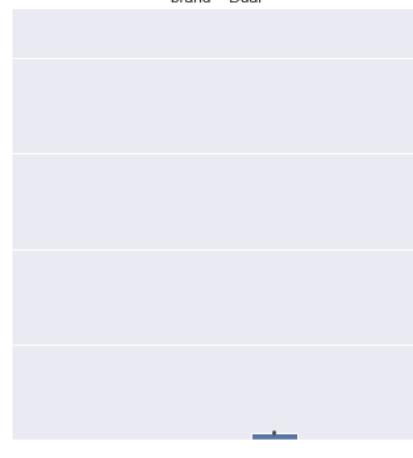
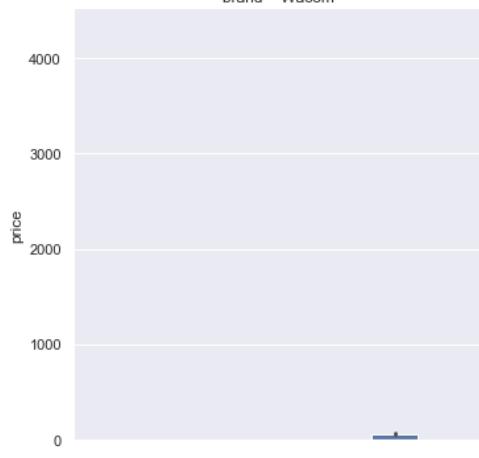
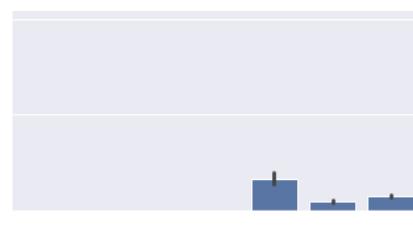


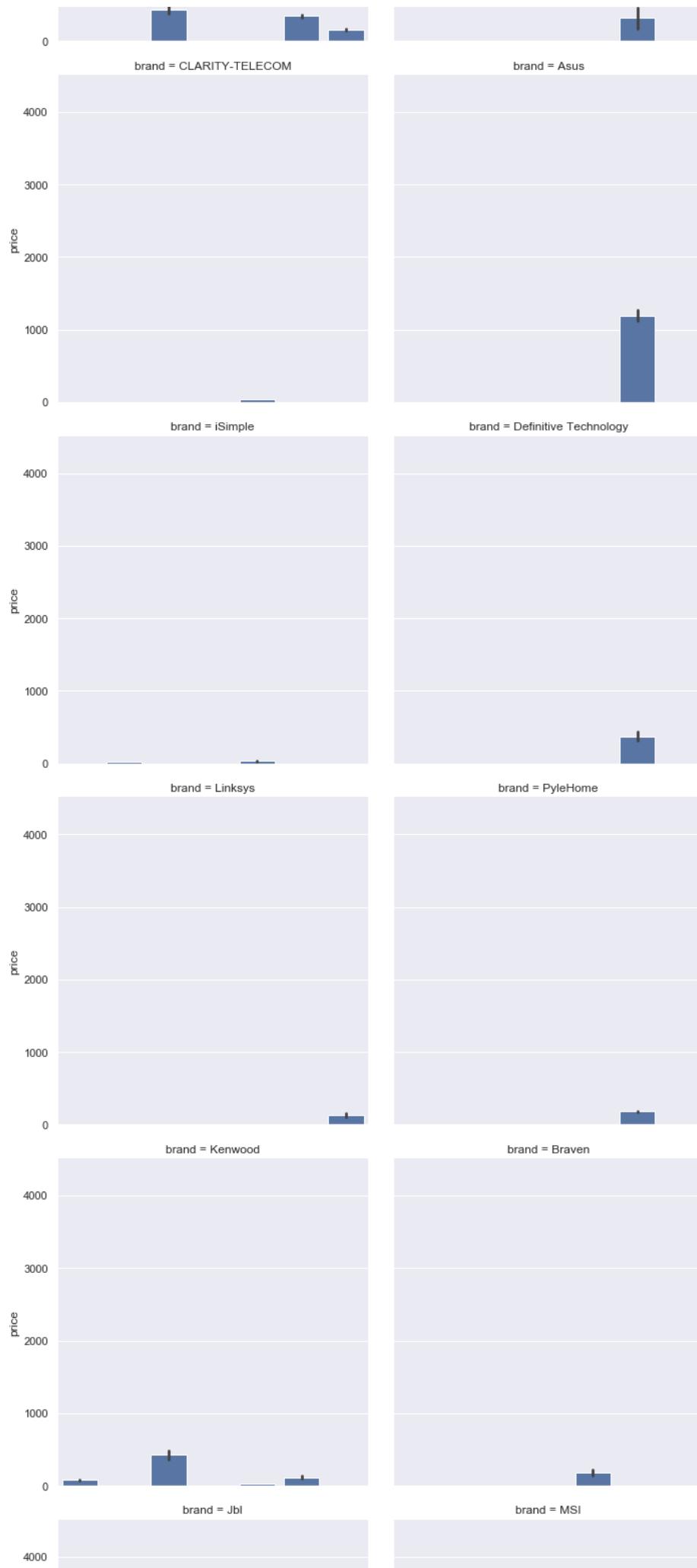


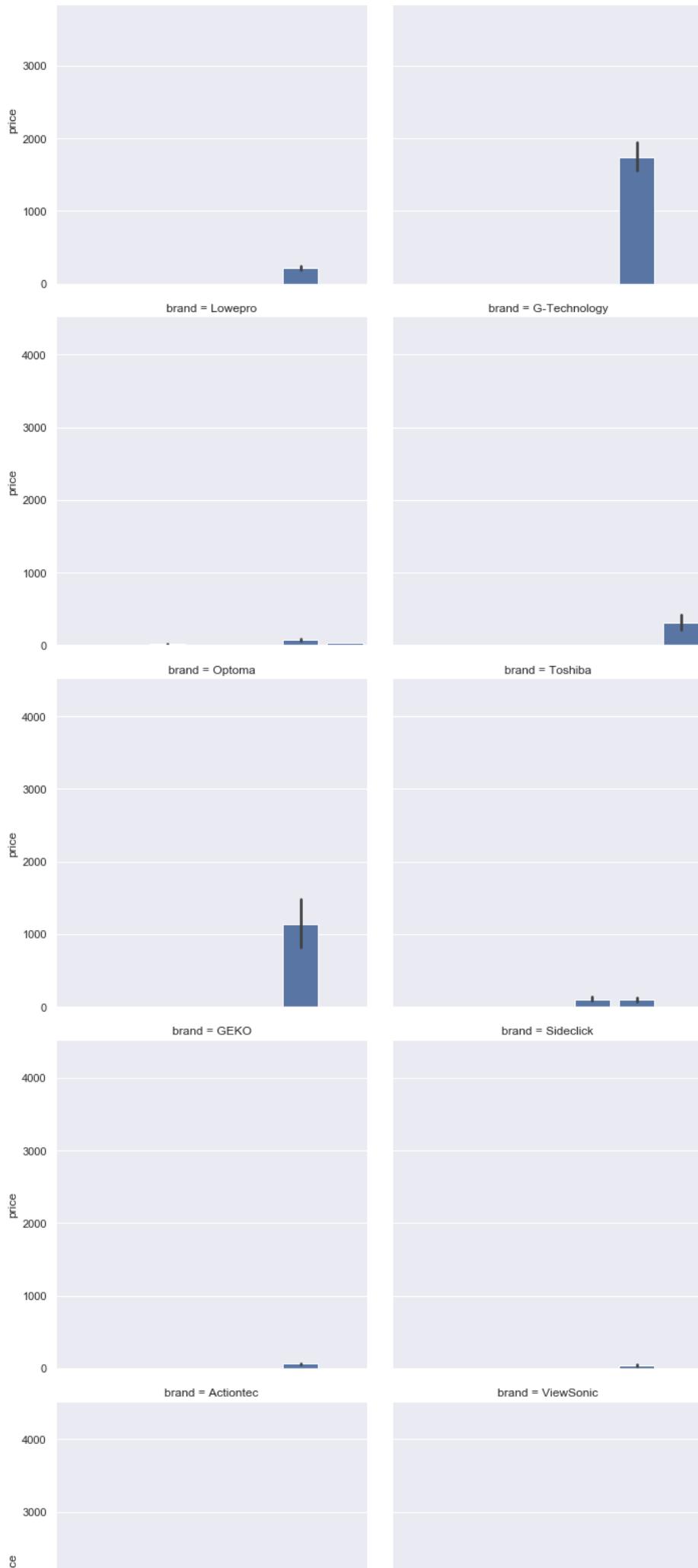


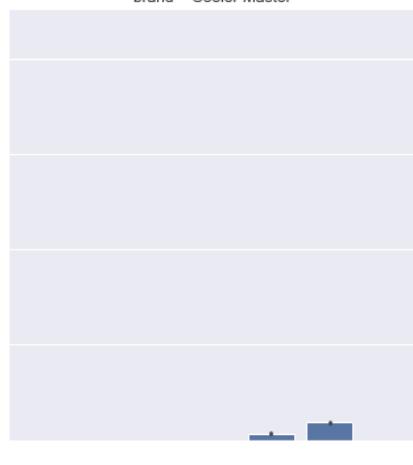
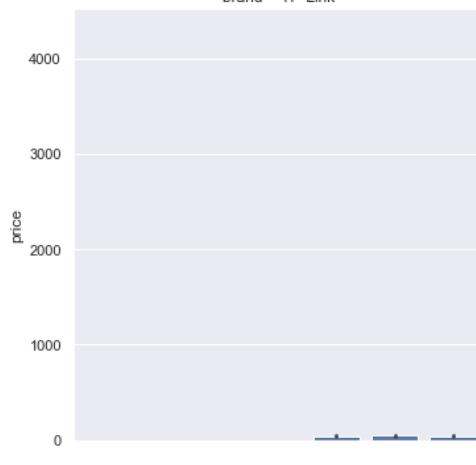
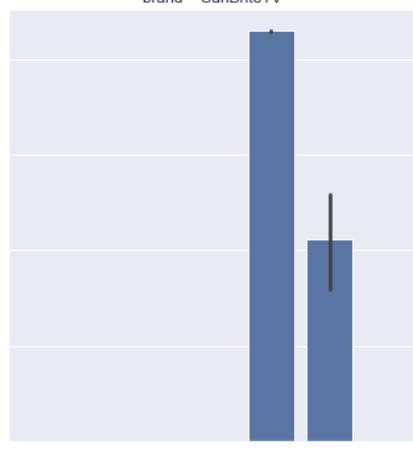
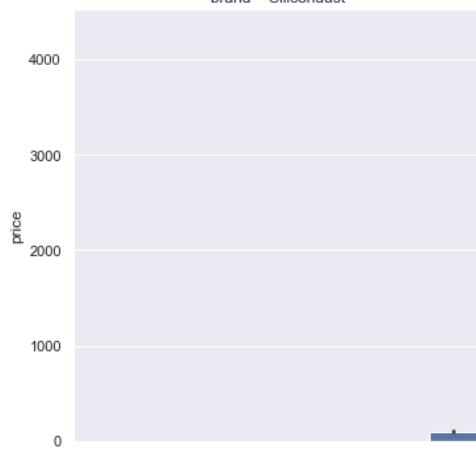
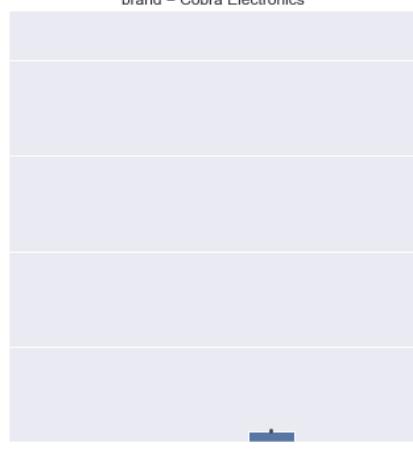
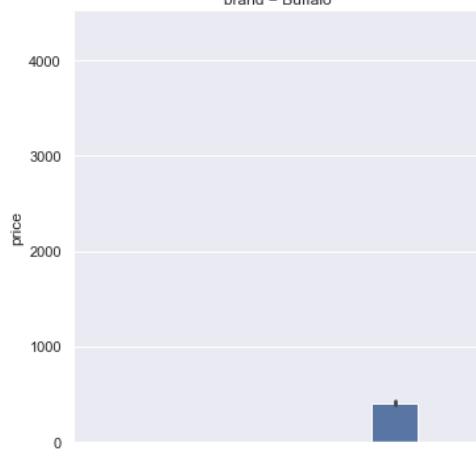
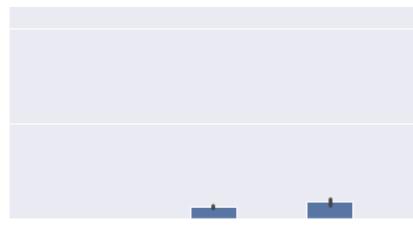


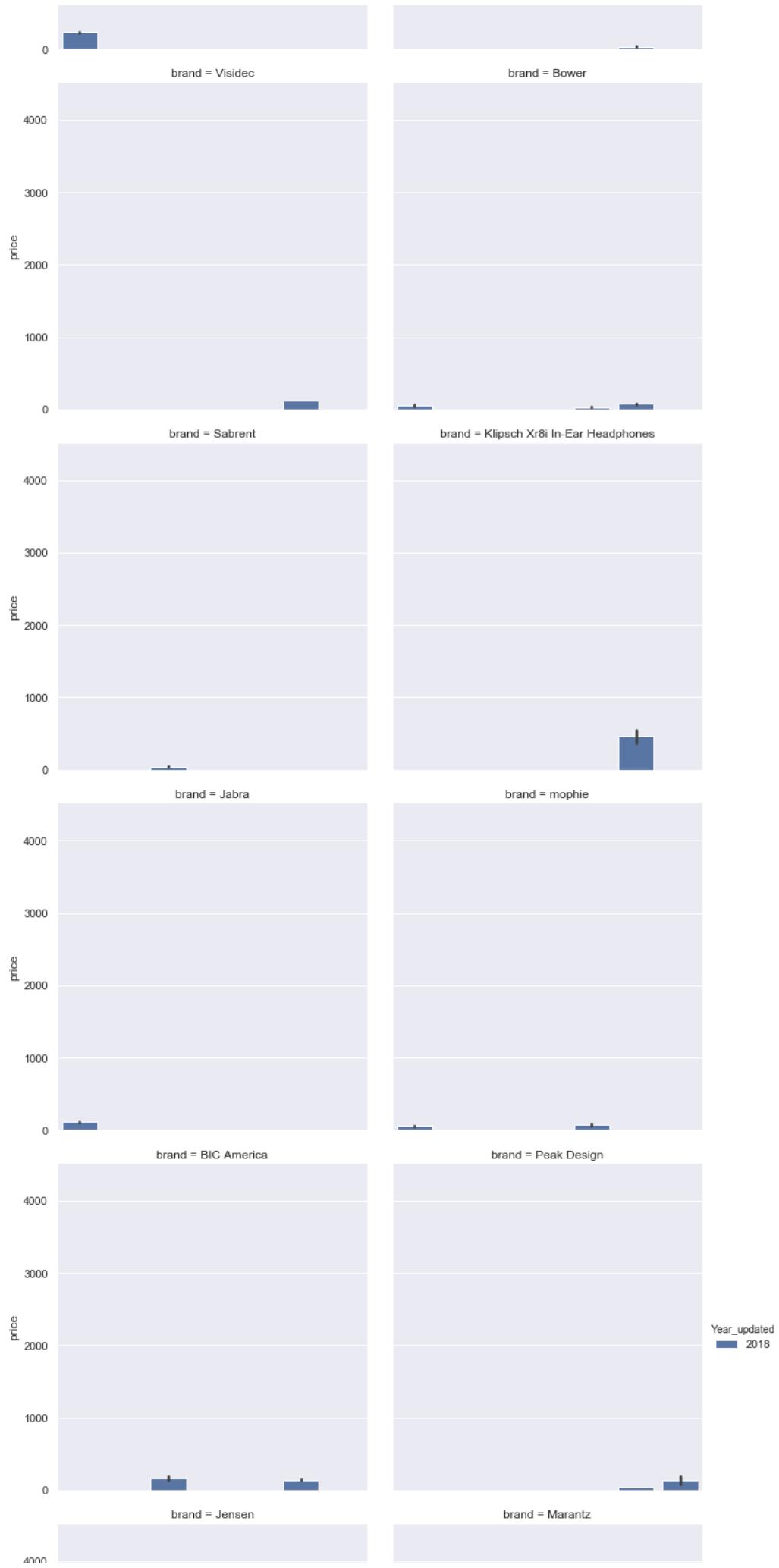


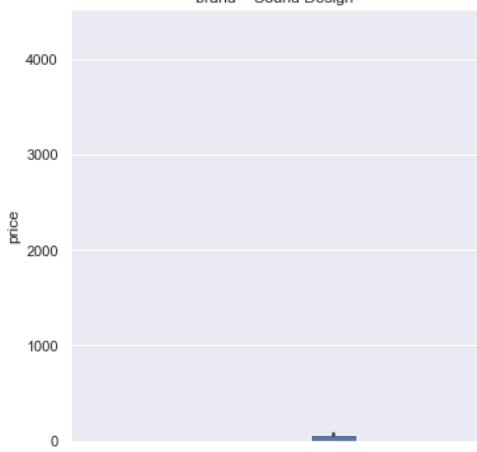
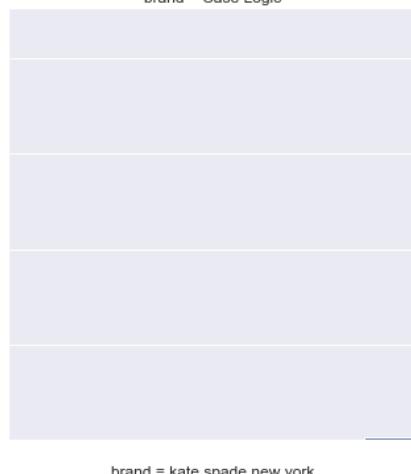
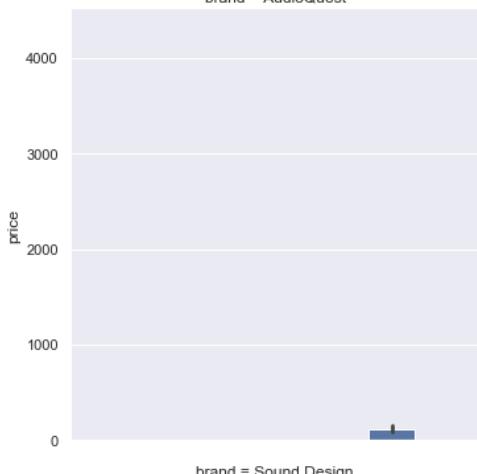
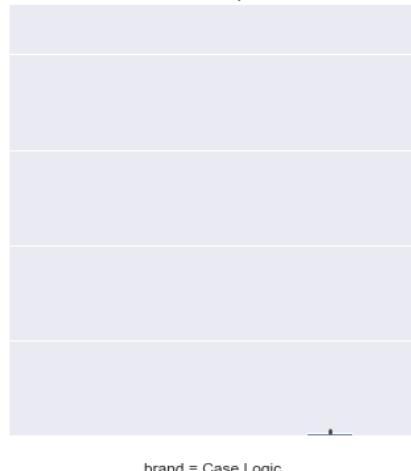
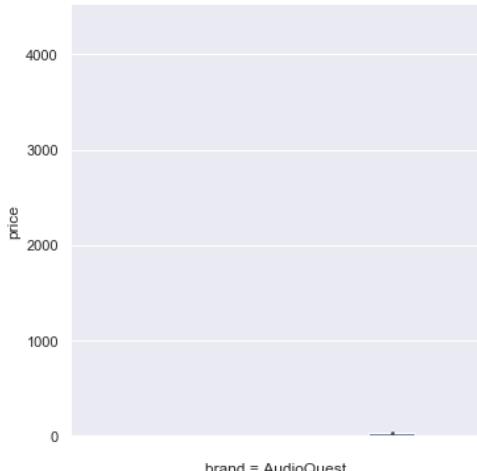
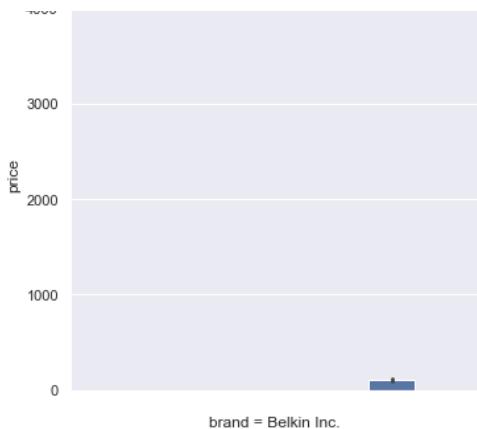


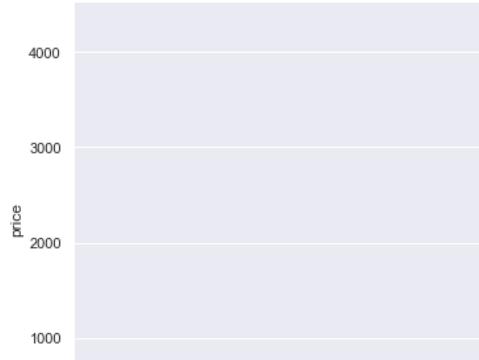
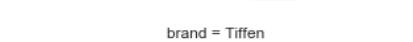
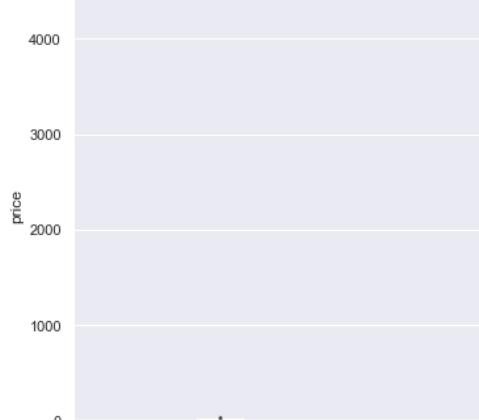
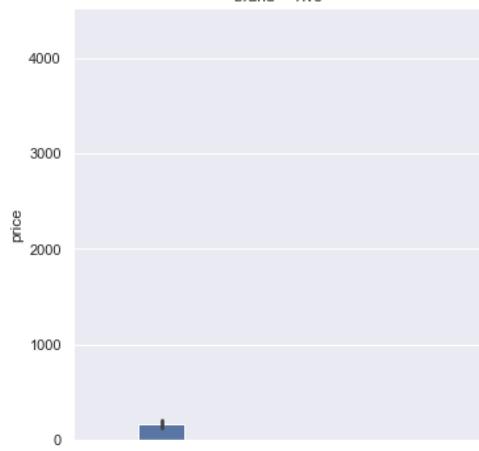
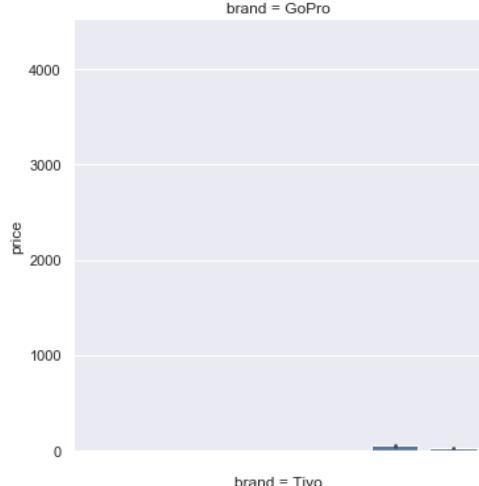
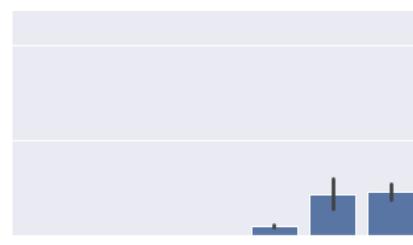


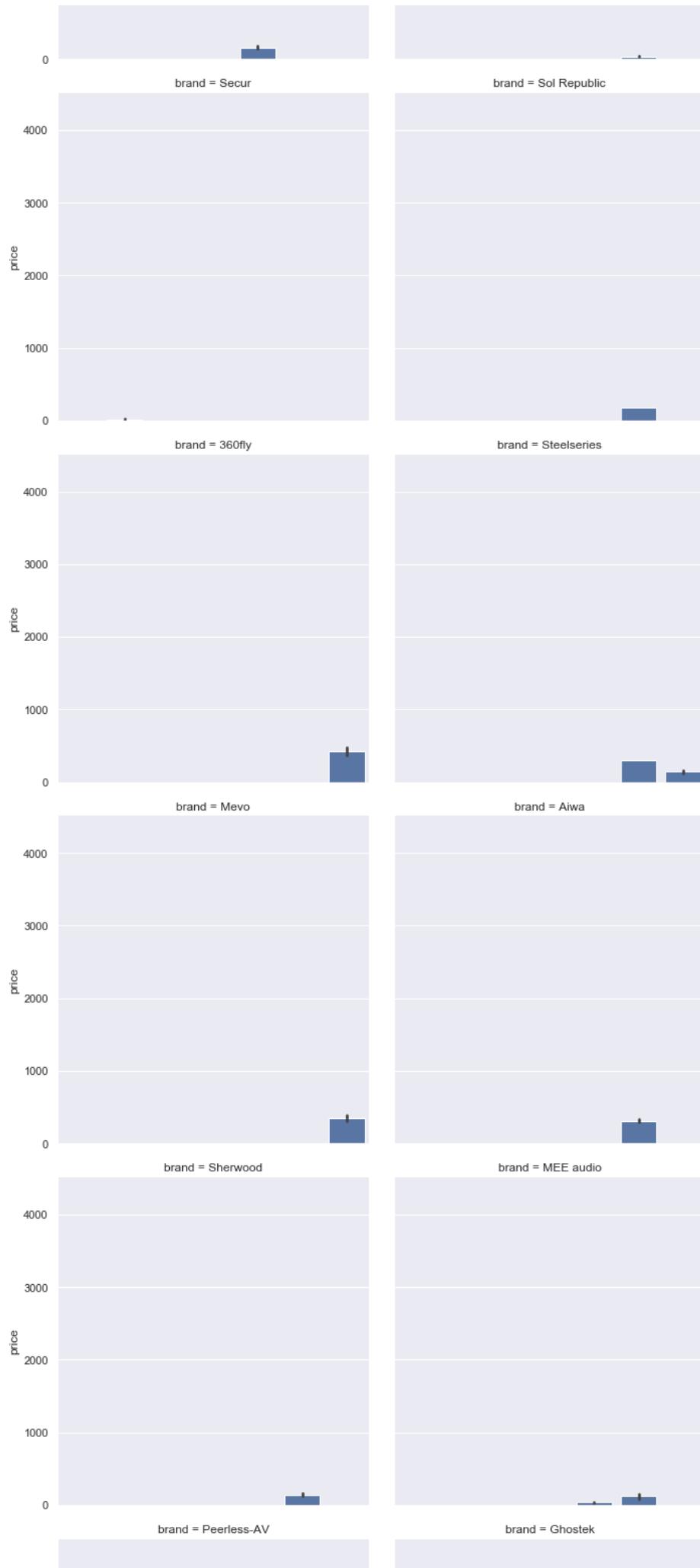


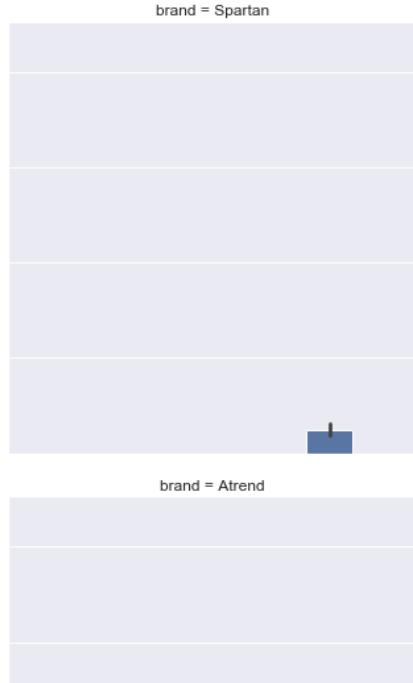
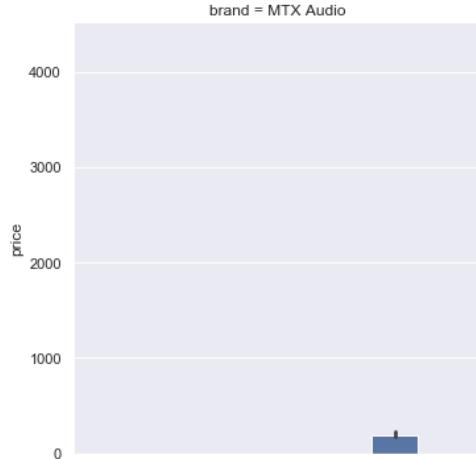
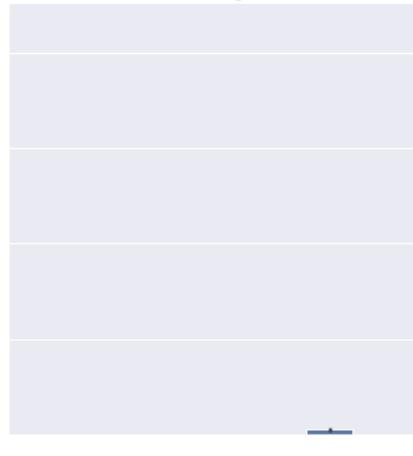
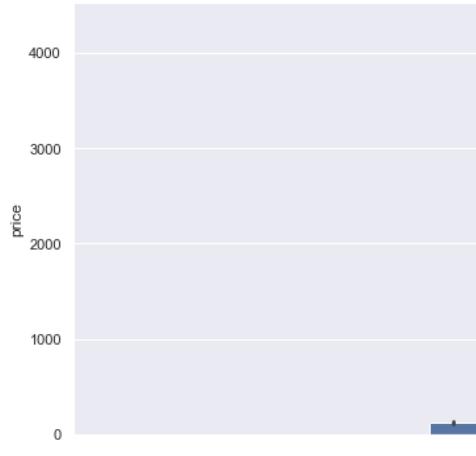
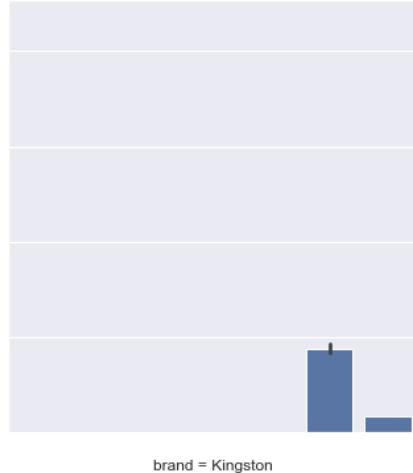
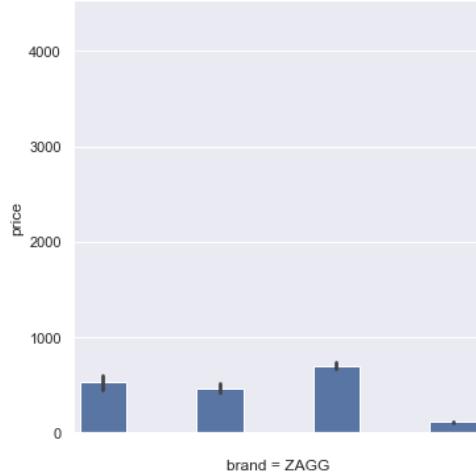
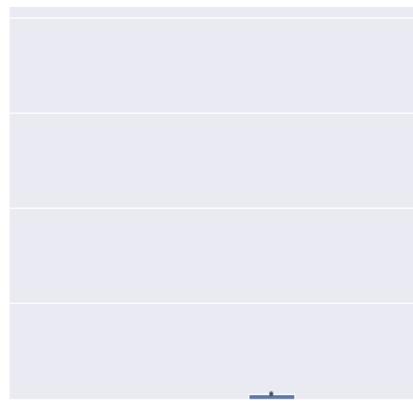
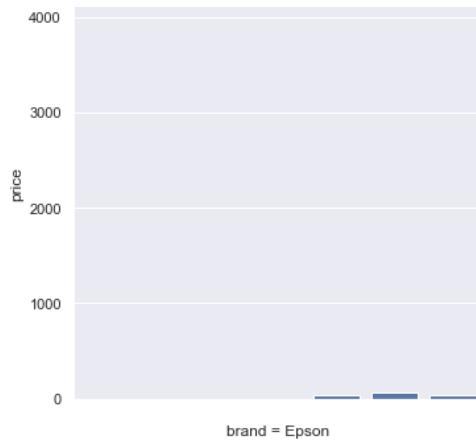


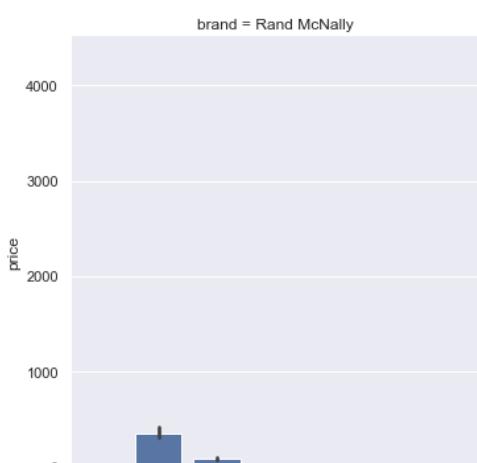
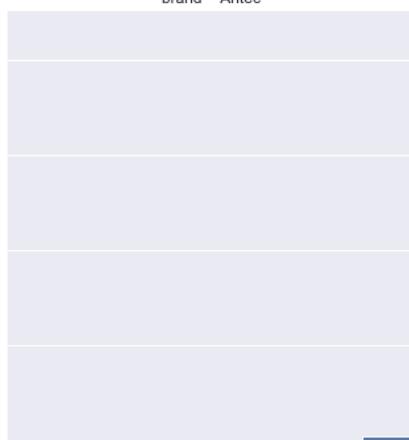
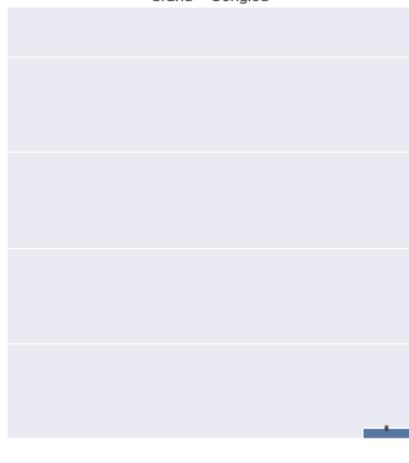
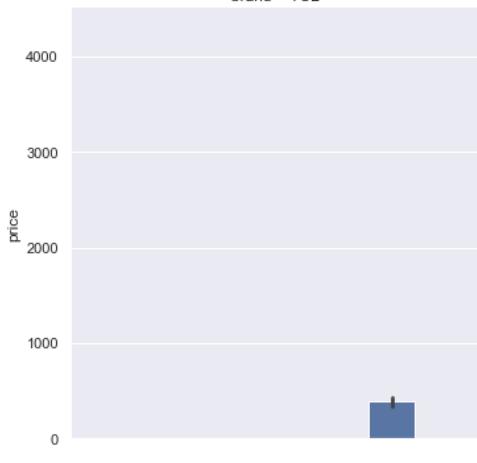


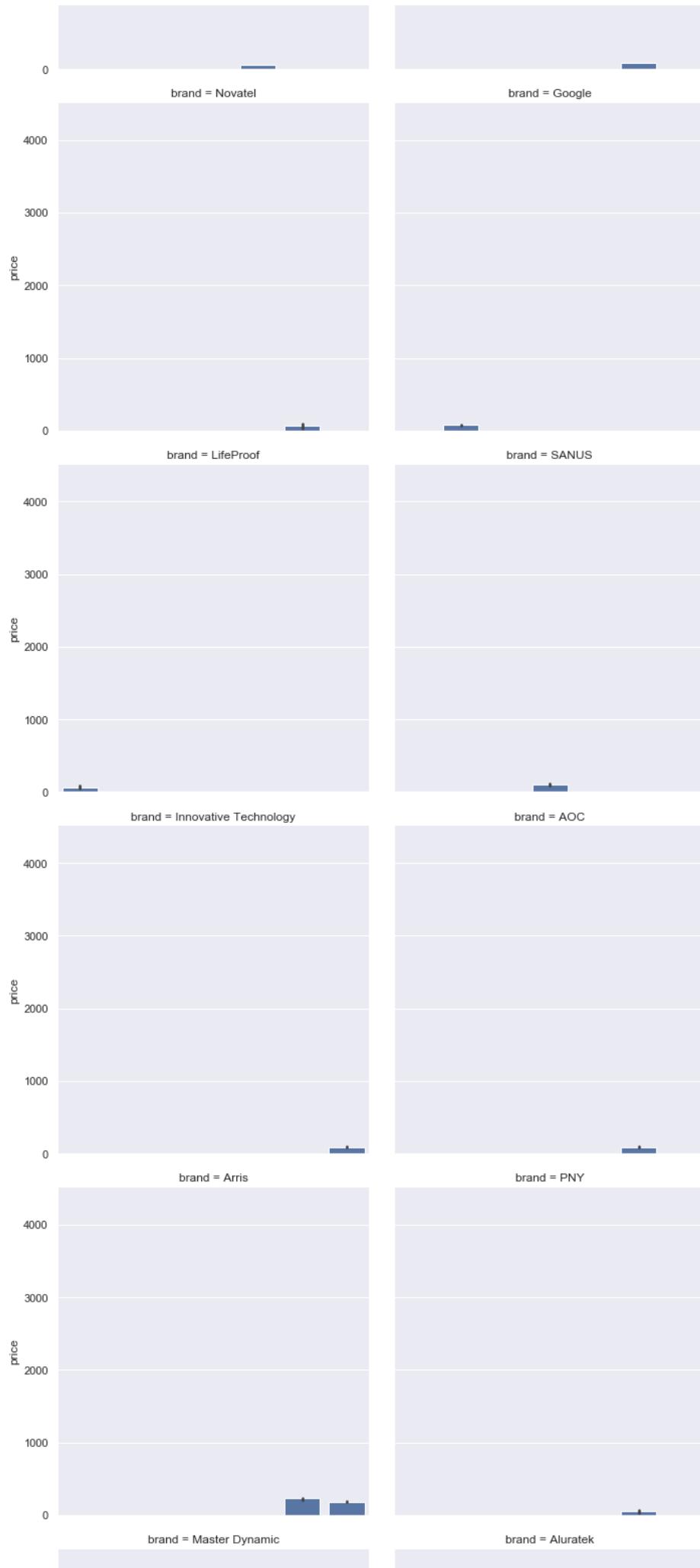


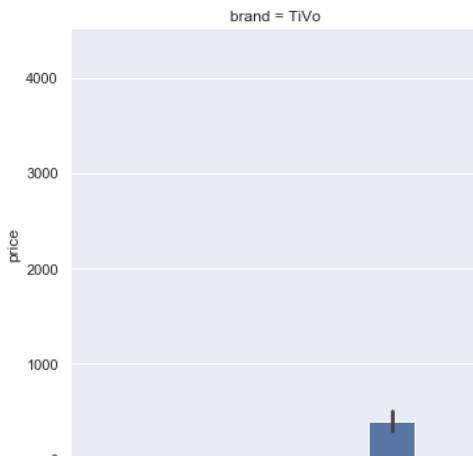
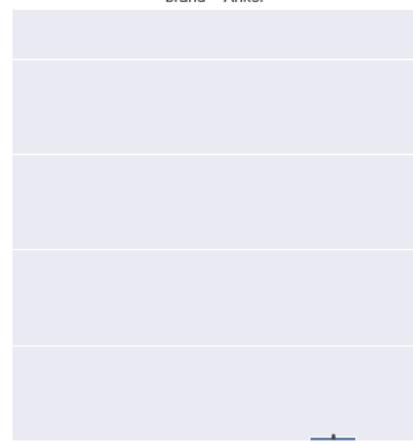
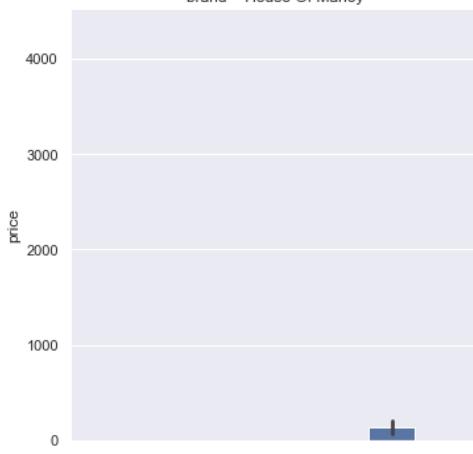
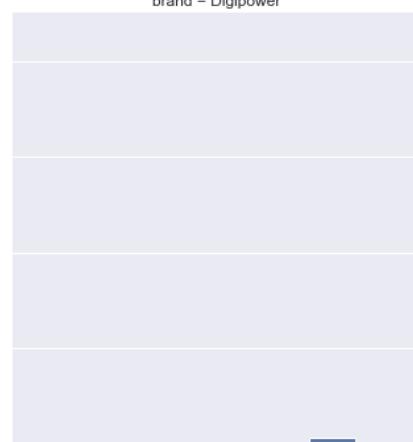
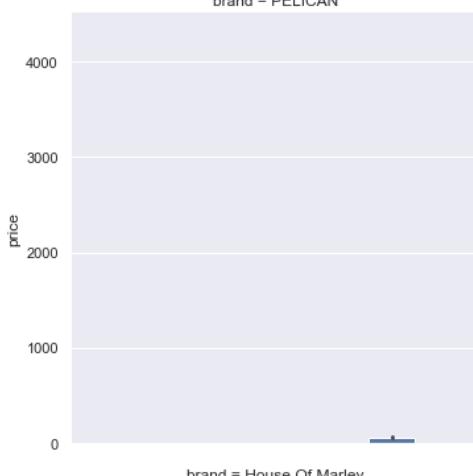
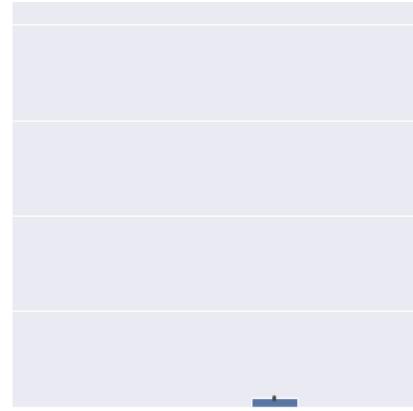
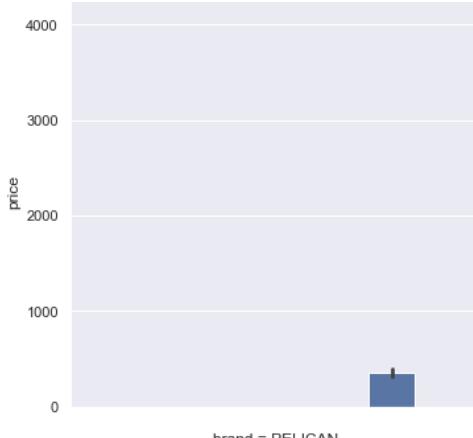


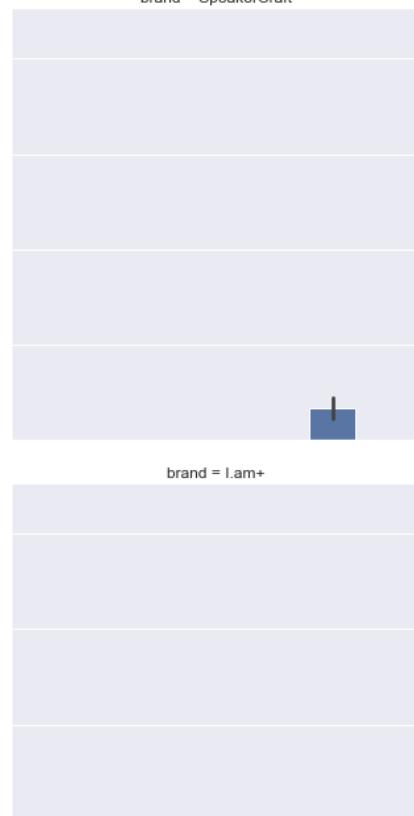
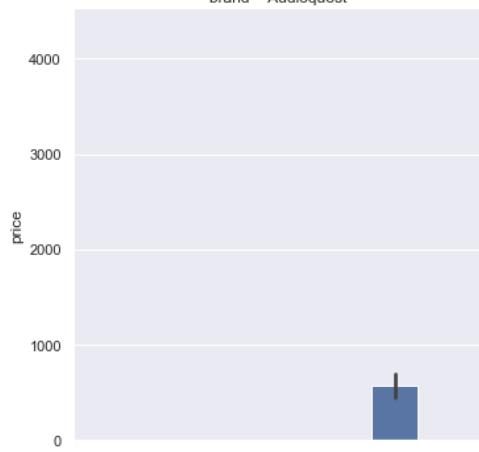
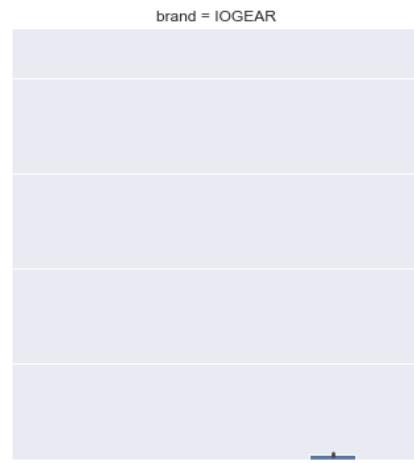
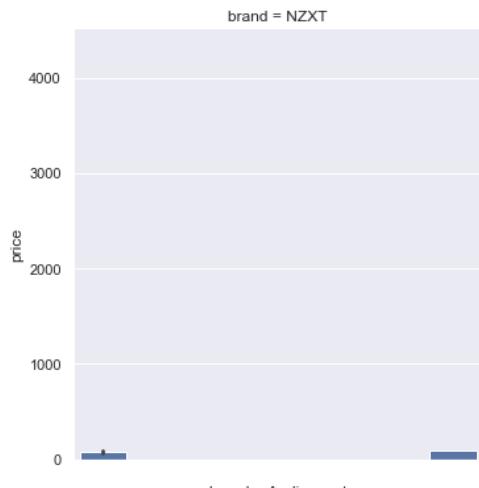
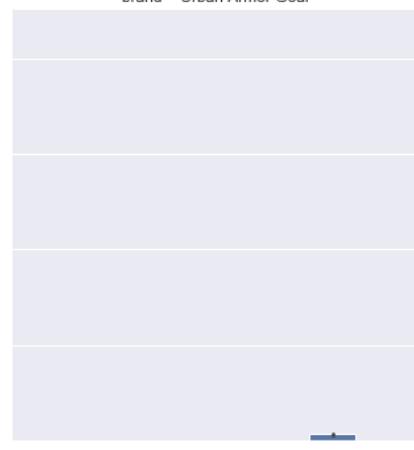
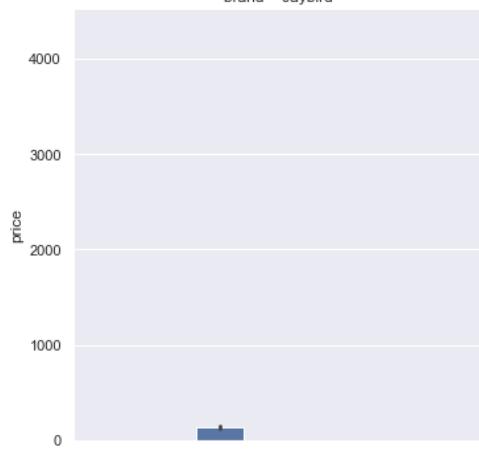
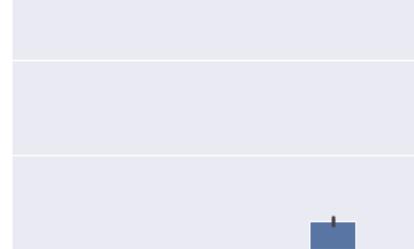
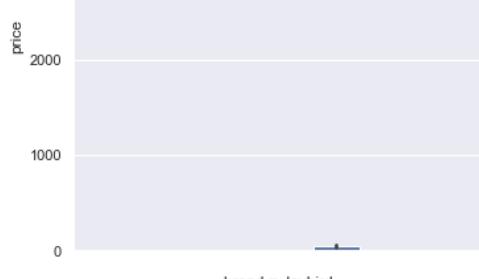




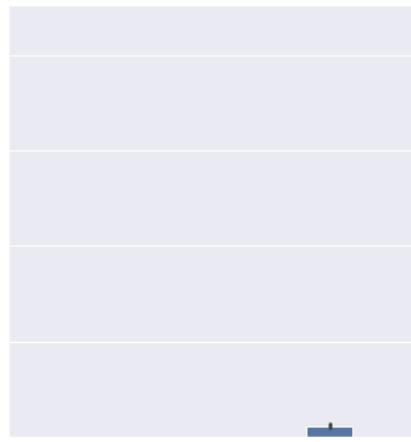
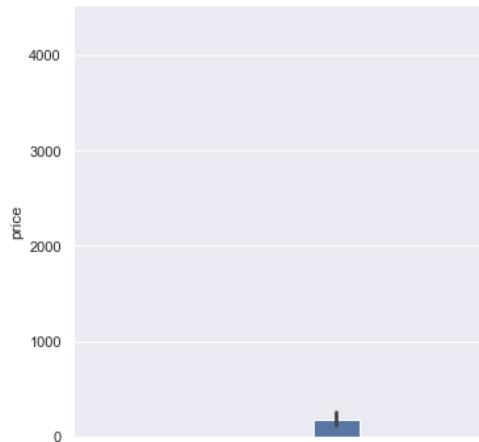
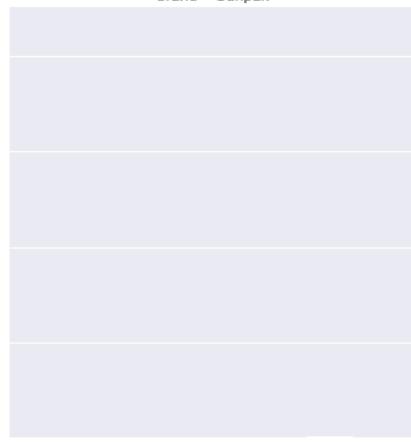
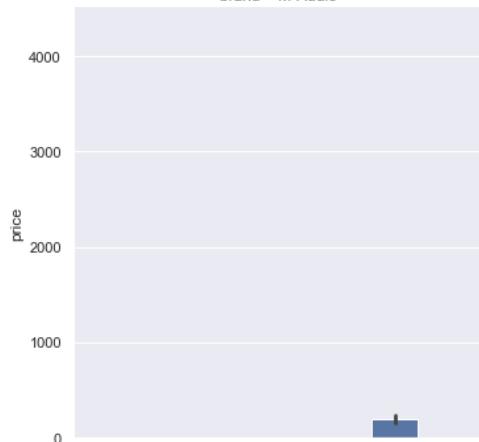
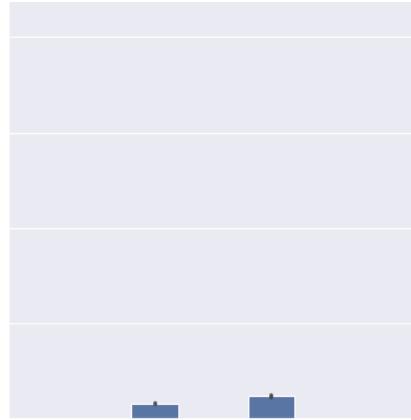
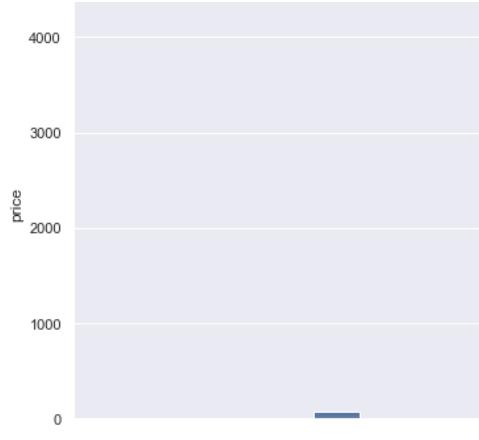


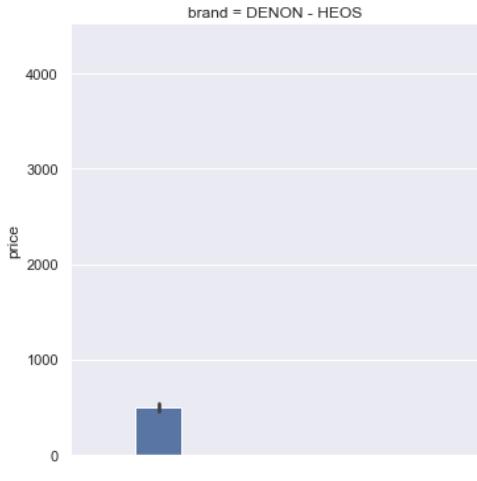
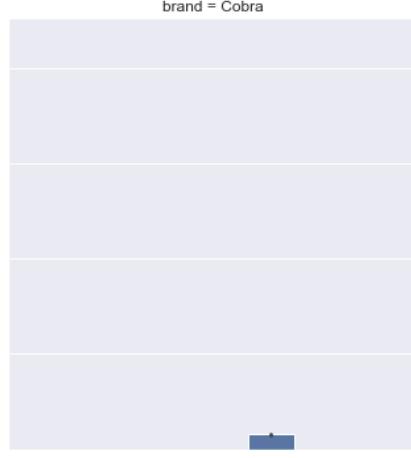
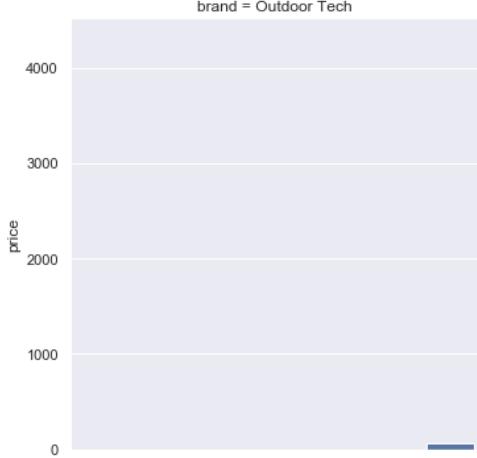
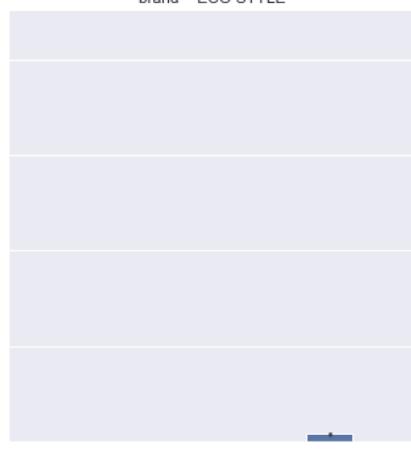
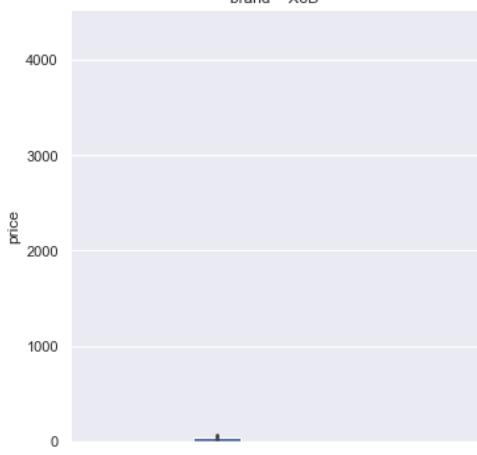


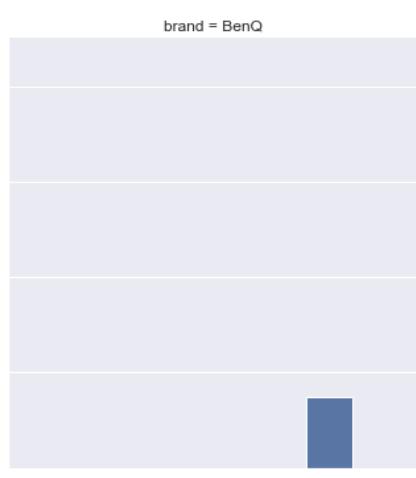
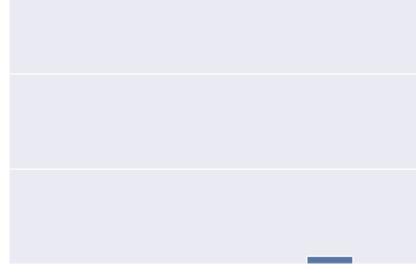
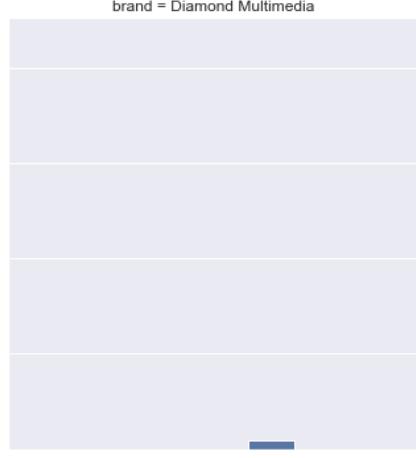
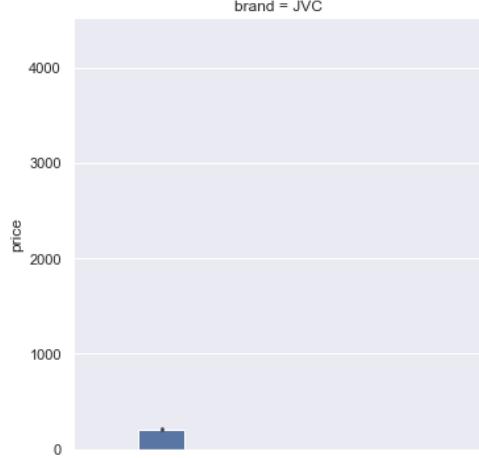


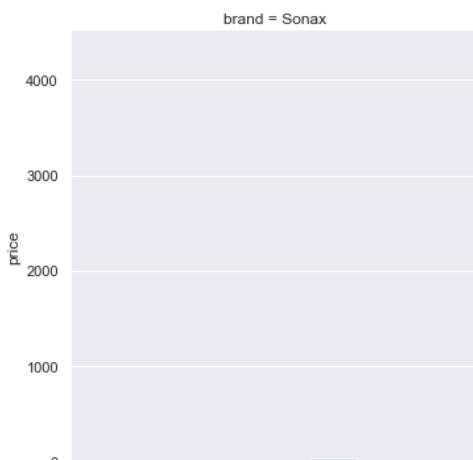
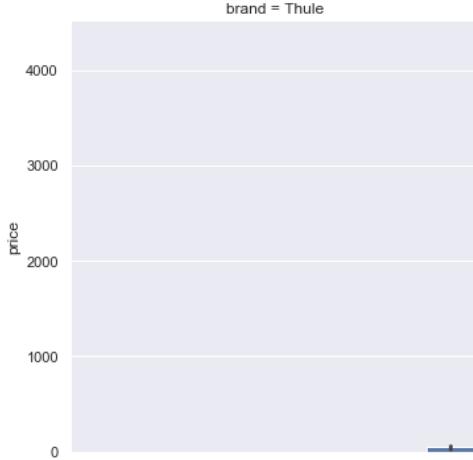
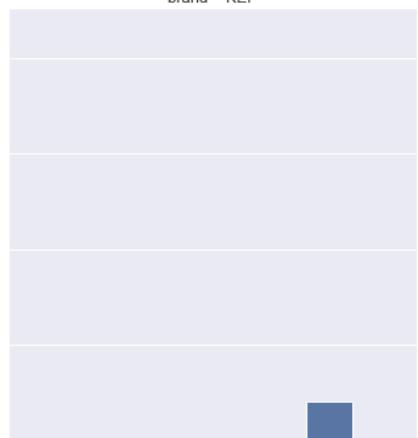
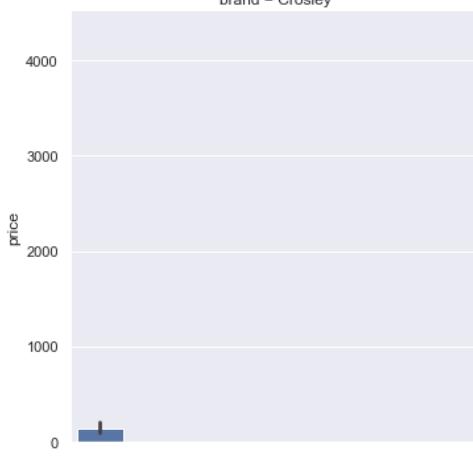
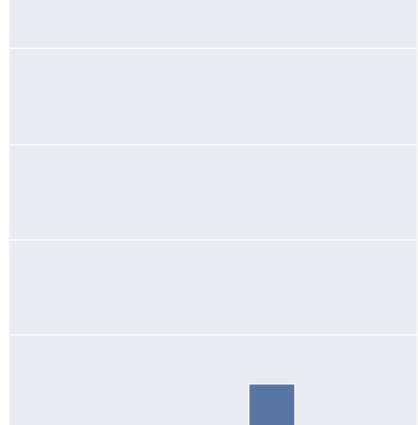
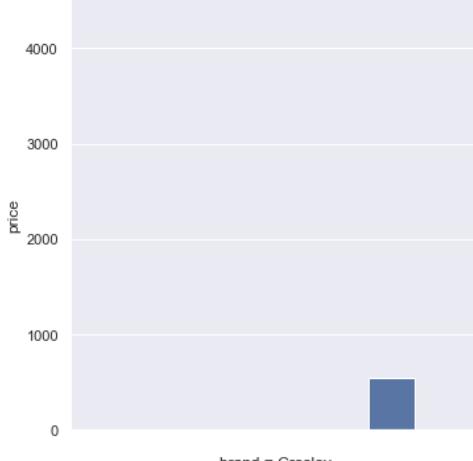


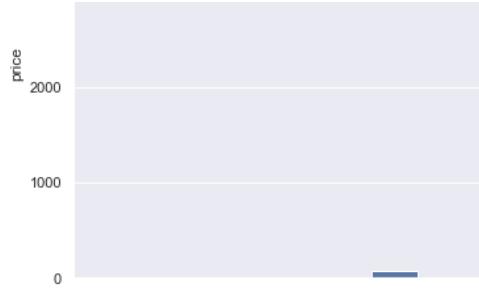




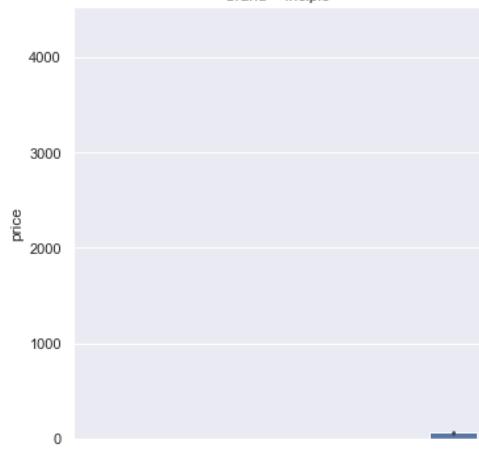
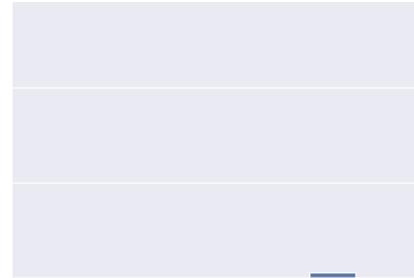




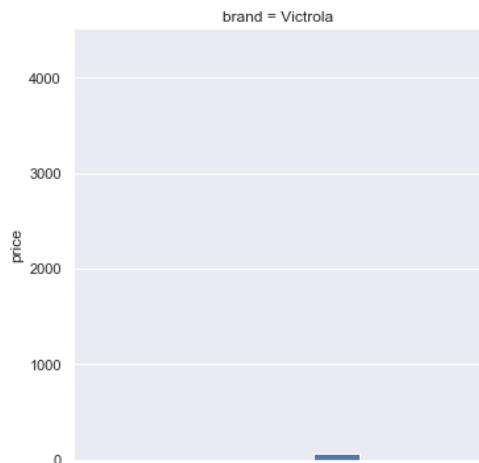
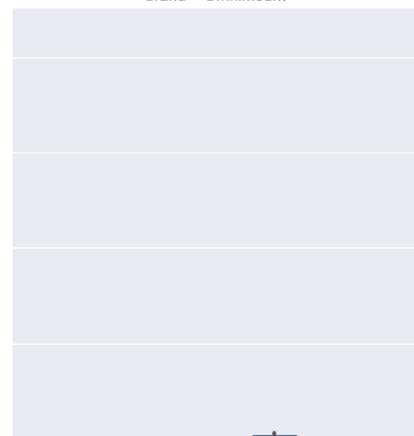




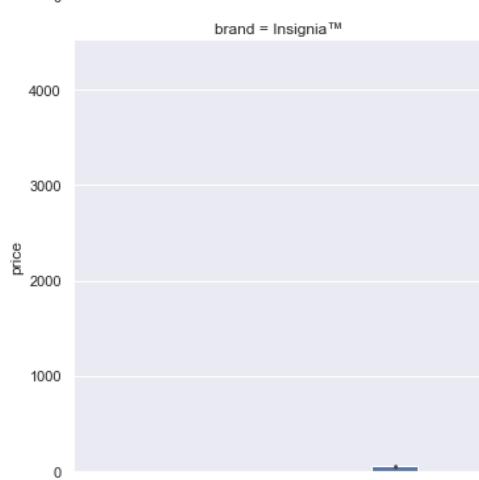
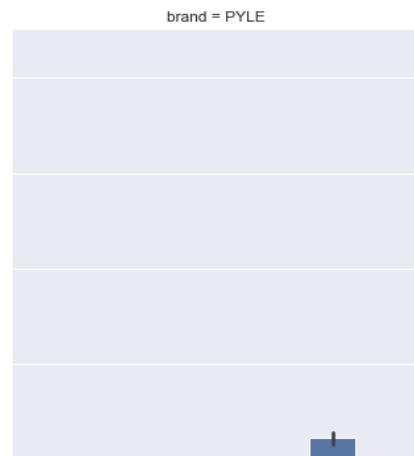
brand = Incipio



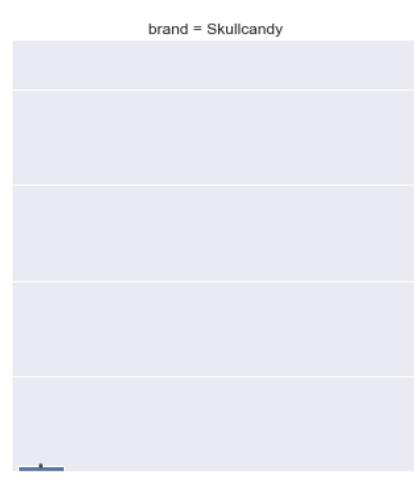
brand = Victrola

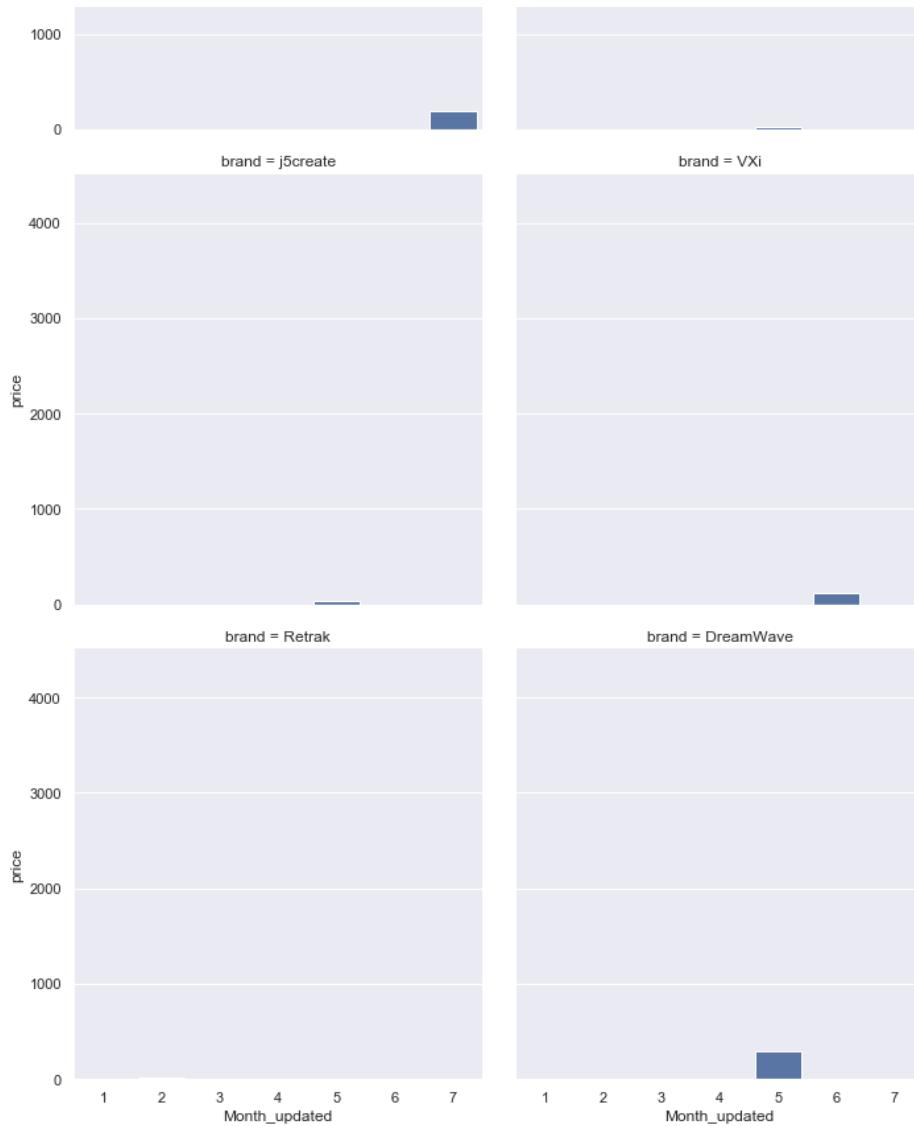


brand = Insignia™



brand = Skybell





Observation

From this graph I can conclude SunBrite TV has more price than other brand
I can also conclude that in 5th and 6th month of 2018 price is highest

For primaryCategories I will apply one-hot encoding

In [75]:

```
import plotly.express as px
import plotly.graph_objects as go
```

In [76]:

```
def histogram(data,path,color,title,xaxis,yaxis):
    fig = px.histogram(data, x=path,color=color)
    fig.update_layout(
        title_text=title,
        xaxis_title_text=xaxis,
```

```
        yaxis_title_text=yaxis,
        bargap=0.2,
        bargroupgap=0.1
    )
fig.show()
```

In [77]:

```
histogram(dt_electronics,"brand", "brand", 'Brand Count', 'Brand', 'Count')
```

In [78]:

```
Number_brand = dt_electronics['brand'].value_counts()
values = Number_brand.values
categories = pd.DataFrame(data=Number_brand.index, columns=["brand"])
categories['values'] = values
```

In [79]:

```
def treemap(categories,title,path,values):
    fig = px.treemap(categories, path=path, values=values, height=700,
                      title=title, color_discrete_sequence = px.colors.sequential
                      .RdBu)
    fig.data[0].textinfo = 'label+text+value'
    fig.show()
```

In [80]:

```
treemap(categories, 'Major Sales of Brand', ['brand'], categories['values'])
```

In [81]:

```
def bar(categories,x,y,color,title,xlab,ylab):
    fig = px.bar(categories, x=x, y=y,
                  color=color,
                  height=400)
    fig.update_layout(
        title_text=title,
        xaxis_title_text=xlab,
        yaxis_title_text=ylab,
        bargap=0.2,
        bargroupgap=0.1
    )
    fig.show()
```

In [82]:

```
bar(categories, categories['brand'][0:10], categories['values'][0:10]
    ,categories['brand'][0:10], 'Top 10 brand', 'brand', 'Count')
```

In [83]:

```
Number_yeardtadded = dt_electronics['Year_dtadd'].value_counts()
years = pd.DataFrame(data=Number_yeardtadded.index, columns=["YEAR"])
years['values'] = Number_yeardtadded.values
```

In [84]:

```
fig = px.pie(years, values='values', names='YEAR', color_discrete_sequence
=px.colors.sequential.RdBu)
fig.show()
```

In [85]:

```
Price_month = dt_electronics['Month_dtadd'].value_counts()  
months = pd.DataFrame(data=Price_month.index, columns=["MONTH"])  
months['values'] = Price_month.values
```

In [86]:

```
fig = px.pie(months, values='values', names='MONTH', color_discrete_sequence=px.colors.sequential.RdBu)  
fig.show()
```

In [87]:

```
| fig = go.Figure(go.Bar(
```

```
x=months[ 'values' ],
y=months[ 'MONTH' ],
marker=dict(
    color='rgb(13,143,129)',
),
    orientation='h'))
fig.update_layout(
    title_text='Price count as per month',
    xaxis_title_text='Count',
    yaxis_title_text='Month',
    bargap=0.2,
    bargroupgap=0.1
)
fig.show()
```

In [88]:

```
histogram(dt_electronics,"brand","brand",'Brand Count','Brand','Count')
```

In [89]:

```
histogram(dt_electronics,"brand","Year_dtadd",'Brand count per Category on  
each Year','Category','Crimes Count on each Year')
```

In [90]:

```
histogram(dt_electronics,"brand","Month_dtadd",'Brand count per Category o  
n each Year','Category','Crimes Count on each Year')
```

In [91]:

```
histogram(dt_electronics,"Year_dtadd","Year_dtadd",'Price Count Yearly Bas  
is','Year','Count')
```

In [92]:

```
high_brand = dt_electronics['brand'].value_counts()  
brand = pd.DataFrame(data=high_brand.index, columns=["Brand"])  
brand['values'] = high_brand.values
```

In [93]:

```
bar(brand,brand['Brand'][0:10],brand['values'][0:10]  
    ,brand['Brand'][0:10],'Top 10 brands','brands',' Price Count')
```

In [94]:

```
dt_electronics.Year_dtadd.unique()
```

Out[94]:

```
array([2015, 2017, 2014, 2016, 2018], dtype=int64)
```

In [95]:

```
dt_2014=dt_electronics[dt_electronics['Year_dtadd']==2014]
```

In [96]:

```
histogram(dt_2014, "brand", "brand", 'Brand count in 2014', 'Brand', 'Count')
```

In [97]:

```
histogram(dt_2014, "brand", "Month_dtadd", 'Brand count per month in 2014', 'Category', 'Brand')
```

In Year 2014 Alpine band has the highest count

In [98]:

```
dt_2015=dt_electronics[dt_electronics['Year_dtadd']==2015]
```

In [99]:

```
histogram(dt_2015,"brand","brand",'Brand count in 2015','Brand','Count')
```

In [100]:

```
histogram(dt_2015, "brand", "Month_dtadd", 'Brand count per month in 2015', 'Category', 'Brand')
```

In Year 2015 Yamaha band has the highest count

In [101]:

```
dt_2016=dt_electronics[dt_electronics['Year_dtadd']==2016]
```

In [102]:

```
histogram(dt_2016, "brand", "brand", 'Brand count in 2016', 'Brand', 'Count')
```

In [103]:

```
histogram(dt_2016, "brand", "Month_dtadd", 'Brand count per month in 2016', 'Category', 'Brand')
```

In Year 2016 Sony band has the highest count

In [104]:

```
dt_2017=dt_electronics[dt_electronics['Year_dtadd']==2017]
```

In [105]:

```
histogram(dt_2017,"brand","brand",'Brand count in 2017','Brand','Count')
```

In [106]:

```
histogram(dt_2017,"brand","Month_dtadd",'Brand count per month in 2017','Category','Brand')
```

In Year 2017 Samsung band has the highest count

In [107]:

```
dt_2018=dt_electronics[dt_electronics['Year_dtadd']==2018]
```

In [108]:

```
histogram(dt_2018,"brand","brand",'Brand count in 2018','Brand','Count')
```

In [109]:

```
histogram(dt_2018, "brand", "Month_dtadd", 'Brand count per month in 2018', 'Category', 'Brand')
```

In Year 2018 Sony band has the highest count

In [110]:

```
dt_electronics.columns
```

Out[110]:

```
Index(['pricesavailability', 'pricescondition', 'pricesdateSeen',
       'pricesmerchant', 'pricesshipping', 'pricessourceURLs', 'asins',
       'brand', 'categories', 'ean', 'imageURLs', 'keys', 'manufacturer',
       'manufacturerNumber', 'name', 'primaryCategories', 'sourceURLs', 'upc',
       'weight', 'price', 'Currency_USD', 'pricesisSale_True',
```

```
'brand_ordinal_labels', 'Month_dtadd', 'Year_dtadd',
'Month_updated',
'Year_updated', 'manufacturer_count'
```

In [111]:

```
primaryCategories=pd.get_dummies(dt_electronics.primaryCategories,drop_fir
st=True,prefix='primaryCategories')
dt_electronics=pd.concat([dt_electronics,primaryCategories],axis=1)
dt_electronics.drop('primaryCategories',axis=1,inplace=True)
dt_electronics.head()
```

Out[111]:

	pricesavailability	pricescondition	pricesdateSeen	pricesmerchant	pricesshipping
0	1590	4645	2017-05-10T20:00:00Z,2017-05-09T15:00:00Z	2109	
1	1590	4645	2017-10-10T02:00:00Z,2017-08-12T03:00:00Z,2017...	2109	
2	1590	4645	2017-10-10T19:00:00Z,2017-09-12T14:00:00Z,2017...	2109	
3	1590	4645	2017-09-08T05:00:00Z,2017-09-18T13:00:00Z,2017...	2109	
4	63	4645	2017-12-05T13:00:00Z	1119	

In [112]:

```
dt_electronics.columns
```

Out[112]:

```
Index(['pricesavailability', 'pricescondition', 'pricesdateSe
en',
       'pricesmerchant', 'pricesshipping', 'pricessourceURL
s', 'asins',
       'brand', 'categories', 'ean', 'imageURLs', 'keys', 'ma
nufacturer',
       'manufacturerNumber', 'name', 'sourceURLs', 'upc', 'we
ight',
       'price',
       'Currency_USD', 'pricesisSale_True', 'brand_ordinal_la
bels',
```

```

'month_ataaa', 'year_ataaa', 'month_upated', 'year_up-
dated',
'manufacturer_count', 'primaryCategories_ Intel Celero-
n',
'primaryCategories_ Siri Eyes Free', 'primaryCategorie-
s_Electronics',
'primaryCategories_Electronics,Furniture'],
dtype='object')

```

I will drop these columns 'categories','brand','ean', 'imageURLs', 'keys', 'manufacturer','manufacturerNumber', 'name', 'sourceURLs', 'upc', 'weight'
 categories data is reflected with primaryCategories so categories can be dropped
 manufacturer column is replaced with manufacturer_count
 brand column is replaced with brand_ordinal_labels
 These columns 'imageURLs', 'keys', 'name', 'sourceURLs', 'upc', 'weight'

In [113]:

```
dt_electronics.drop(['categories', 'brand', 'ean', 'imageURLs', 'keys', 'man-
ufacturer', 'manufacturerNumber', 'name', 'sourceURLs', 'upc', 'weight'],ax-
is=1,inplace=True)
```

In [114]:

```
dt_electronics.drop(['manufacturer_count'],axis=1,inplace=True)
```

Feature Selection

I will apply different techniques

1. Corelation

In [115]:

```
dt_electronics.corr()
```

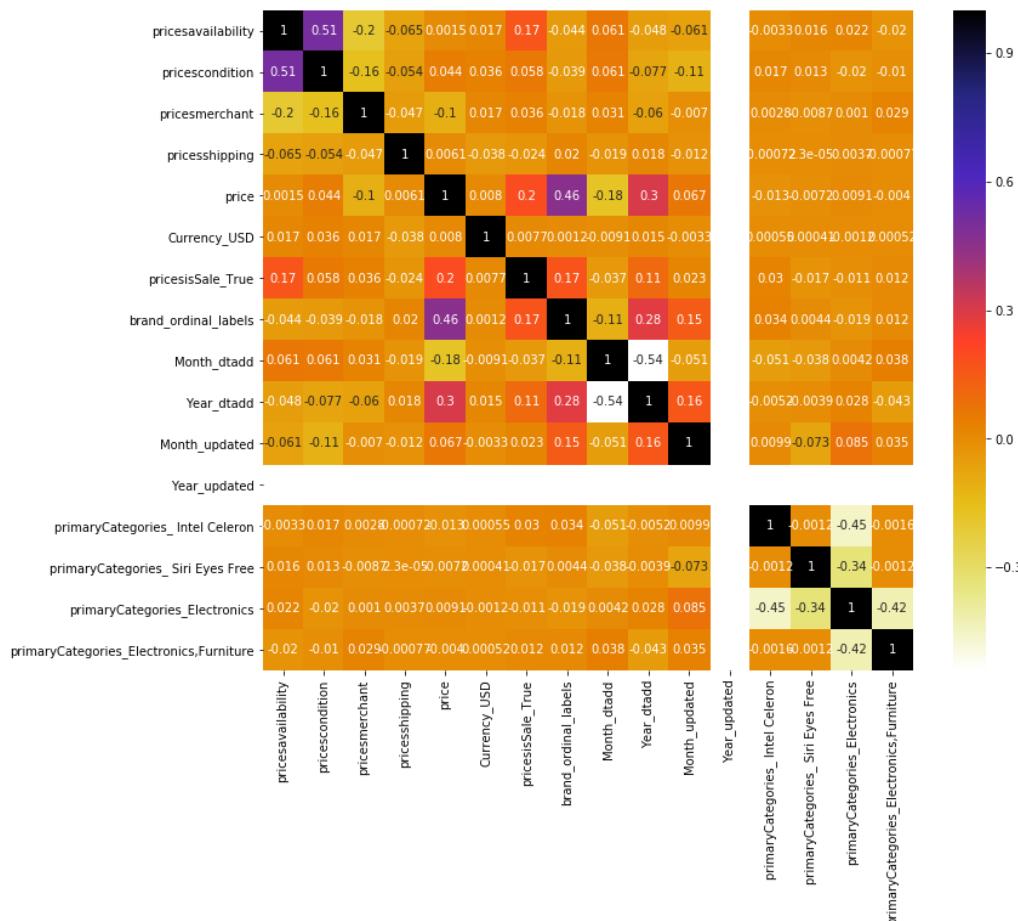
Out[115]:

	pricesavailability	pricescondition	pricesmer-
pricesavailability	1.000000	0.508854	-0.19
pricescondition	0.508854	1.000000	-0.19
pricesmerchant	-0.199715	-0.156314	1.00
pricesshipping	-0.065426	-0.054341	-0.04
price	0.001523	0.044039	-0.10
Currency_USD	0.017189	0.035772	0.01
pricesisSale_True	0.165858	0.057858	0.01

	brand_ordinal_labels	-0.043529	-0.038755	-0.0
	Month_dtadd	0.061236	0.061123	0.0:
	Year_dtadd	-0.048377	-0.077155	-0.0
	Month_updated	-0.061485	-0.113257	-0.0
	Year_updated	NaN	NaN	
	primaryCategories_Intel Celeron	-0.003307	0.016780	0.00
	primaryCategories_Siri Eyes Free	0.016289	0.012502	-0.0
	primaryCategories_Electronics	0.021845	-0.019811	0.00
	primaryCategories_Electronics,Furniture	-0.020494	-0.010420	0.0:

In [116]:

```
plt.figure(figsize=(12,10))
cor = dt_electronics.corr()
sns.heatmap(cor, annot=True, cmap=plt.cm.CMRmap_r)
plt.show()
```



In [117]:

```
def correlation_feature(dataset, threshold):
    col_corr = set() # Set of all the names of correlated columns
    corr_matrix = dataset.corr()
    for i in range(len(corr_matrix.columns)):
        for j in range(i):
            if abs(corr_matrix.iloc[i, j]) > threshold: # we are interested in absolute coeff value
                col_corr.add(corr_matrix.columns[j])
```

```
        colname = corr_matrix.columns[i] # getting the name of column
        col_corr.add(colname)
    return col_corr
```

Setting the threshold value to 0.5

In [118]:

```
corr_features = correlation_feature(dt_electronics, 0.5)
len(set(corr_features))
```

Out[118]:

2

In [119]:

```
corr_features
```

Out[119]:

```
{'Year_dtadd', 'pricescondition'}
```

2. Finding zero variance columns

In [120]:

```
zero_var=dt_electronics.var()[dt_electronics.var()==0].index.values
```

In [121]:

```
zero_var
```

Out[121]:

```
array(['Year_updated'], dtype=object)
```

Observation

From the above Feature Selection process I can conclude that I can drop these three columns 'Year_dtadd', 'pricescondition', 'Year_updated'

In [122]:

```
dt_electronics.columns
```

Out[122]:

```
Index(['pricesavailability', 'pricescondition', 'pricesdateSeen',
       'pricesmerchant', 'pricesshipping', 'pricessourceURLs',
       'asins',
       'brand', 'Currency_HSD', 'InvoicingSale_Third', 'Brand_o
```

```
        price , currency_usd , pricesdateSeen=True , train_u
rdinal_labels',
        'Month_dtadd' , 'Year_dtadd' , 'Month_updated' , 'Year_up
dated',
        'primaryCategories_ Intel Celeron' , 'primaryCategories
_Siri Eyes Free',
        'primaryCategories_Electronics',
        'primaryCategories_Electronics,Furniture'],
dtype='object')
```

In [123]:

```
dt_electronics.drop('pricesdateSeen',axis=1,inplace=True)
```

In [124]:

```
dt_electronics.drop(['pricessourceURLs','asins'],axis=1,inplace=True)
```

In [125]:

```
dt_electronics.drop(['Year_dtadd','pricescondition','Year_updated'],axis=1
,inplace=True)
```

Model Development

In this section I will use differnt machine learning algorithm to predict the sales for these stores

Performance Metric

I am going to use R square value as performance metric. The intution behind R square value is when R square is closer to 1 indicates that it's a better model

In [126]:

```
x=dt_electronics.drop('price',axis=1)
y=dt_electronics['price']
```

Splitting the dataset into train and test dataset

In [127]:

```
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3,random_st
ate=42)
```

In [128]:

```
x_train.shape,y_train.shape
```

Out[128]:

```
((3805, 12), (3805,))
```

Linear Regression

In [129]:

```
model_lr=LinearRegression()
model_lr.fit(x_train,y_train)
model_pred_tr=model_lr.predict(x_train)
model_pred_test=model_lr.predict(x_test)
```

In [130]:

```
print('R2 square for train',r2_score(y_train,model_pred_tr))
print('R2 square for test',r2_score(y_test,model_pred_test))
```

R2 square for train 0.25029309674174804
R2 square for test 0.25411079403472836

Observation

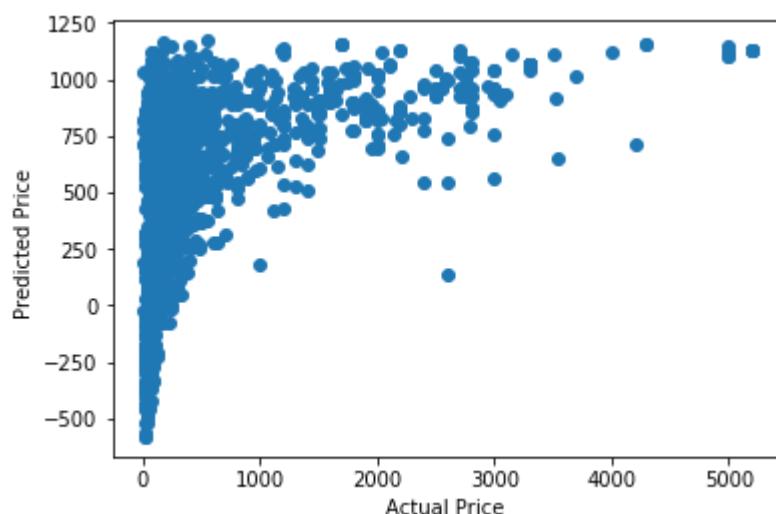
As I know if R square value is close to 1 it indicate that it's a good model
From the above R square value for train and test I can see that both are not close to 1
For both train and test set my model is giving same value for R Square which indicate that model is not over fit

In [131]:

```
plt.scatter(y_test,model_pred_test)
plt.xlabel('Actual Price ')
plt.ylabel('Predicted Price ')
```

Out[131]:

Text(0, 0.5, 'Predicted Price ')



Random Forest

In [132]:

```
model_rand=RandomForestRegressor()  
model_rand.fit(x_train,y_train)  
model_rand_train=model_rand.predict(x_train)  
model_rand_test=model_rand.predict(x_test)
```

C:\Users\Subhasish Das\Anaconda3\lib\site-packages\sklearn\ensemble\forest.py:245: FutureWarning:

The default value of n_estimators will change from 10 in version 0.20 to 100 in 0.22.

In [133]:

```
print('R2 square for train',r2_score(y_train,model_rand_train))  
print('R2 square for test',r2_score(y_test,model_rand_test))
```

R2 square for train 0.8613360320273933
R2 square for test 0.642558658070556

Observation

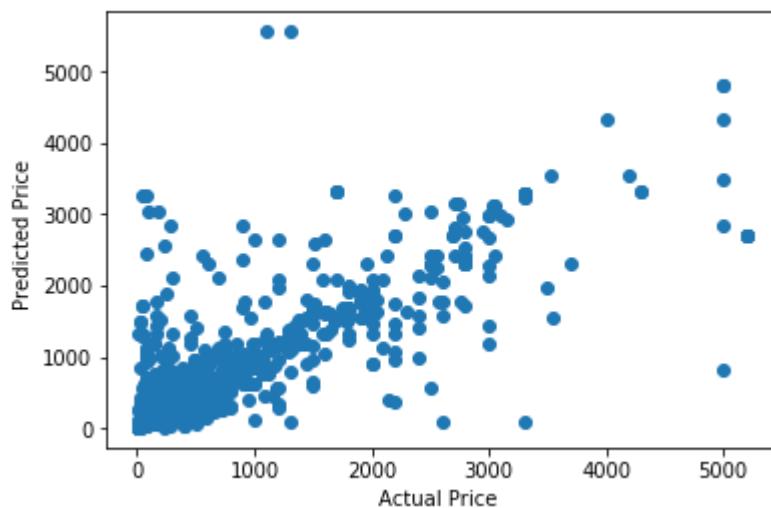
From Random forest model I am getting R2 square value more closer to 1 which indicate that this model is better than Linear Regression model

In [134]:

```
plt.scatter(y_test,model_rand_test)  
plt.xlabel('Actual Price ')  
plt.ylabel('Predicted Price ')
```

Out[134]:

Text(0, 0.5, 'Predicted Price ')



Decision Tree

In [135]:

```
dct_mod=DecisionTreeRegressor()  
dct_mod.fit(x_train,y_train)  
dct_mod_train=dct_mod.predict(x_train)  
dct_mod_test=dct_mod.predict(x_test)
```

In [136]:

```
print('R2 square for train',r2_score(y_train,dct_mod_train))  
print('R2 square for test',r2_score(y_test,dct_mod_test))
```

R2 square for train 0.887933143003788
R2 square for test 0.580190414306418

Observation

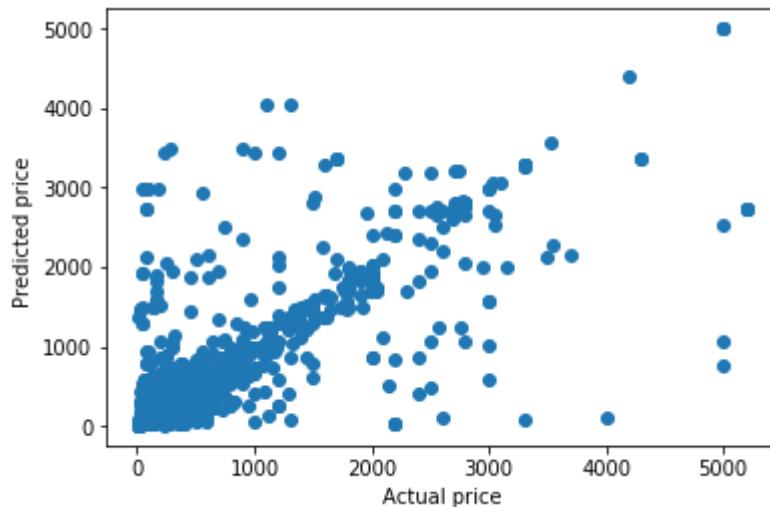
From Decision Tree Regressor model I am getting R2 square value .92 which is less than the value I got from Random Forest

In [137]:

```
plt.scatter(y_test,dct_mod_test)  
plt.xlabel('Actual price')  
plt.ylabel('Predicted price ')
```

Out[137]:

Text(0, 0.5, 'Predicted price ')



Conclusion

This Problem is not about Prediction. This problem is about the the price distribution of brands and categories on a yearly basis

That distribution is showed on above graph that which band which manufacturer and which category has highest price over a period of 3 years

In []:

In []: