

# **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

### **Importing All the library**

```
In [2]: 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
```

```
C:\Users\Subhasish Das\Anaconda3\lib\site-packages\dask\config.py:168: YAMLLoadWarning: calling yaml.load() without Loader=... is deprecated, as the default Loader is unsafe. Please read https://msg.pyyaml.org/load for full details.
    data = yaml.load(f.read()) or {}
C:\Users\Subhasish Das\Anaconda3\lib\site-packages\distributed\config.py:20: YAMLLoadWarning: calling yaml.load() without Loader=... is deprecated, as the default Loader is unsafe. Please read https://msg.pyyaml.org/load for full details.
    defaults = yaml.load(f)
```

```
In [3]: pd.set_option('display.max_rows', 100000)
pd.set_option('display.max_columns',500000)
```

## Loading Data Set

```
In [4]: dt_electronics=pd.read_csv("electronics_products_pricing.csv")
```

```
In [5]: dt_electronics.head()
```

Out[5]:

		id	prices.availability	prices.condition	prices.currency	price
0		AVphrugr1cnluZ0-FOeH	Yes	New	USD	10T2 0
1		AVrl6FDbv8e3D1O-lm4R	Yes	New	USD	10TC 12T03
2		AVpiLlubilAPnD_xBoTa	Yes	New	USD	10T1 12T14
3		AVpgQP5vLJeJML43LQbd	Yes	New	USD	08TC 18T13
4		AV1YDsмоGV-KLJ3adcbe	More on the Way	New	USD	

```
In [6]: dt_electronics.shape
```

Out[6]: (5436, 25)

```
In [7]: dt_electronics.describe()
```

Out[7]:

	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 [8]: dt_electronics.isnull().sum()
```

```
Out[8]: 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 [9]: dt_electronics.columns
```

```
Out[9]: Index(['id', 'prices.availability', 'prices.condition', 'prices.currency',  
              'prices.dateSeen', 'prices.isOnSale', '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 [10]: 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.s
                                     ourceURLs':'pricessourceURLs'},inplace=True)
```

```
In [11]: dt_electronics.head()
```

Out[11]:

			id	pricesavailability	pricescondition	pricescurrency	price
0			AVphrugrlcnluZ0-FOeH	Yes	New	USD	10T20: 05-
1			AVrl6FDbv8e3D1O-lm4R	Yes	New	USD	10T02: 12T03:0
2			AVpiLlubilAPnD_xBoTa	Yes	New	USD	10T19: 12T14:0
3			AVpgQP5vLJeJML43LQbd	Yes	New	USD	08T05: 18T13:0
4			AV1YDsmoGV-KLJ3adcbe	More on the Way	New	USD	(

I will drop the Id column

```
In [12]: dt_electronics.drop('id',axis=1,inplace=True)
```

Finding different unique values for each columns

```
In [13]: 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 [14]: dt_electronics.dtypes
```

```
Out[14]: 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 [15]: dt_electronics.pricesavailability.value_counts()
```

```
Out[15]: 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  
Retired         1  
FALSE          1  
7 available      1  
Name: pricesavailability, dtype: int64
```

```
In [16]: dt_electronics.pricesavailability.unique()
```

```
Out[16]: array(['Yes', 'More on the Way', 'In Stock', 'yes', 'TRUE',  
   'Special Order', 'Out Of Stock', 'undefined', 'No', 'sold',  
   '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 [17]: pricesavailability_map=dt_electronics['pricesavailability'].value_counts()  
().to_dict()
```

```
In [18]: pricesavailability_map
```

```
Out[18]: {'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,  
          'Retired': 1,  
          'FALSE': 1,  
          '7 available': 1}
```

here I will replace 'In Stock' with 2395

```
In [19]: dt_electronics['pricesavailability']=dt_electronics['pricesavailability']
          .map(pricesavailability_map)
          dt_electronics.head()
```

Out[19]:

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

Now I will apply the same process for pricescondition

```
In [20]: dt_electronics.pricescondition.unique()
```

```
Out[20]: 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 L
   oom Tubing, Depth: 6.5" (165mm) (top) 11.2" (285mm) (bottom), Item Weigh
   t: 18.5 lbs., Frequency Response 25Hz - 500Hz, Line Output, Max Power: 11
   00 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 1
   0", 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 su
   rround, Nominal Impedance 4 ohm, Rugged fiberboard construction with thi
   ck carpet covering, Warranty: 1 Year Manufacturer Warranty, MOSFET Power,
   Weight: 6 lbs, Width: 17.7" (450mm), Condition: Brand New!, Low Pass Cros
   sover, List item, RMS Power: 250 Watt x 1 Channel @ 4 Ohm, Remote Bass Co
   ntrol 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 Bla
   ck RCA, SPL (db @ 1w/1m) 88dB, New Dual SBX101 10" 500 Watts Car Audio Su
   bwoofer Sub + Ported Sub Enclosure, (20) 4" Wire Ties',
   'refurbished',
   'New Kicker BT2 41IK5BT2V2 Wireless Bluetooth USB Audio System Bla
   ck + Remote, Power Supply (volts, ampere): 24, 2.9, Square Reflex Subwoof
   er (in., mm): 6 x 6", Stereo Amp Power with DSP (watts): 50, App for cust
   omizing - 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!, Tw
   eeter (in., cm): 3/4"'],
   dtype=object)
```

```
In [21]: dt_electronics.pricescondition.value_counts()
```

```
Out[21]: New  
4645  
new  
537  
Used  
118  
New other (see details)  
63  
Manufacturer refurbished  
43  
Seller refurbished  
15  
Refurbished  
8  
pre-owned  
3  
refurbished  
2  
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: 1 2.8" (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 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  
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): 8.8 7" 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  
Name: pricescondition, dtype: int64
```

```
In [22]: pricescondition_map=dt_electronics['pricescondition'].value_counts().to_dict()
pricescondition_map
```

```
Out[22]: {'New': 4645,
 'new': 537,
 'Used': 118,
 'New other (see details)': 63,
 'Manufacturer refurbished': 43,
 'Seller refurbished': 15,
 'Refurbished': 8,
 'pre-owned': 3,
 'refurbished': 2,
 '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 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,
 '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}
```

```
In [23]: dt_electronics['pricescondition']=dt_electronics['pricescondition'].map(pricescondition_map)
dt_electronics.head()
```

Out[23]:

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

For pricescurrency and pricesisSale I will apply one-hot encoding

```
In [24]: dt_electronics.pricescurrency.unique()
```

Out[24]: array(['USD', 'CAD'], dtype=object)

```
In [25]: dt_electronics.pricesisSale.unique()
```

Out[25]: array([False, True])

```
In [26]: Currency=pd.get_dummies(dt_electronics.pricescurrency,drop_first=True, prefix='Currency')
```

```
In [27]: dt_electronics=pd.concat([dt_electronics,Currency],axis=1)
```

```
In [28]: dt_electronics.drop('pricescurrency',axis=1,inplace=True)
```

```
In [29]: 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[29]:

	pricesavailability	pricescondition	pricesdateSeen	pricesmerchant	pricesht
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 mos o

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

```
In [30]: dt_electronics.pricesmerchant.unique()
```

```
Out[30]: 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', 'newopenbo
x',
   'Hyper Microsystems, Inc.', 'Incredible Cellular LLC',
   'Wayfair - Bestbuy.com', 'OneStopShop', 'OneCall', 'tbdeals',
   'gadget-town', 'bestserviceguys', 'goldstar_tech', 'JB Tool Sale
s',
   'platinummicro', 'SharpPrices', 'Adorama Camera',
   'GRAMOPHONE (Authorized Dealer)', 'antonline', 'photovideo4less',
   'AUDIO WATT STORE', 'bargainbrokers', 'jbtoolsales02',
   'electronic_express', '305mia65', 'weoutfit', 'ELAVA LLC',
   'www-sonicelectronix-com', 'closeoutmartdeals', 'Dell.com',
   'Soltech US Corp', 'CosmeticMall', 'raghav_205', 'mpe007',
   'wirelessalliance', 'redtagcamera', 'best_buy', 'techhouse520',
   'kays_place_too', 'Focus Camera - Walmart.com', 'bling',
   'ravencat3', 'TEKONLINE', 'Malelo and Company', 'UnbeatableSale',
   'autosounds_warehouse', 'HappEshopper', 'Hot Deals 4 Less?',
   'LC Bargain Sales', 'refurbforless', 'E-Warehouse-Deals',
   'geekdeal', 'zoneusa', 'outletmp3', 'gmatexpert',
   'Unbeatable Deals', 'mamastore_co', 'echo-and-optics',
   'Monoprice Inc', 'Tech For Less Inc', 'Aiwa Inc.', 'playalegend',
   'singledadsunited', 'hsum8885', 'DemProductSales', 'nps',
   'All in Retail', 'BuyDirect & Save', 'Adorama', 'HPP Enterprises',
   "Samy's Camera", 'sando_distributions', 'Global Audio Video LLC',
   'lucky-31067', 'The Pixel Hub', 'Newegg', 'DLR Hunter',
   'proaudio654', 'Wayfair - Walmart.com', 'Olivabel',
   'Pro AV Dealer', 'new-technologies', 'Made Technology', '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', 'ipadelectronics',
```

'Fab4Electronics', 'listenup', 'istrawberrize', 'aSavings',  
'Acumen Disc', 'pro\_car\_audio\_consulting', 'Googol Shop',  
'antonline.com', 'Audio Advice', 'datu503', 'outletpc',  
'a1Components', 'DigitalKynect', 'robertscamera', 'RackGo',  
'soundoftristate', 'poppy9780', 'Nexit, Inc.',  
'True Modern Electronics', 'amc0052010', 'extremenetwork',  
'compuceil.techn', 'BuilderDepot, Inc.', 'OliveElect, LLC',  
'Music123 - Walmart.com', 'Platinum Micro', 'iphoneguy19',  
'AV Toy Store', 'Zeek and Zellies', 'ryajoh7', 'Audiosavings',  
'truemodernelectronicsusa', 'HONESTDEALS', 'djwoodie', 'Music123',  
'jbcuttlefish', 'chribrackma-0', 'bornewinner', 'umpquablue',  
'raffaffle', 'ejuky', 'cashinwhi', '4surpluscity', 'tech1store',  
'autosounds4less', 'jazzman2275', 'rideausales',  
'1 Stop Electronics Center, Inc.', 'dieg5651\_8xrznj2zi',  
'chrismadron148', 'Always In Touch', 'beachaudio', 'tunnyllc',  
'PT Etech', '6ave', 'efesales', 'shopcelldeals', 'HookedOnTronic  
s',  
'eCopUSA', 'iTech INTL', 'BidDeal', 'Lightning Deals',  
'ONE STOP SHOP', 'Give 5 To Cancer, Inc', 'Mian Brothers Inc',  
'Tri State Camera', 'VMInnovations', 'gearbrokers', 'mygoods',  
'Cameta Camera', 'HiFiSoundConnection', 'Xcess Limited', 'axxii',  
'aimusa', 'zal-digital', 'topusedcomputers', 'g4mmlj', 'Yogi Com  
p',  
'a4c', 'felixg7488', 'Audiolab Stereo & Video Center', 'lite\_cor  
p',  
'Electronics101deals', 'cynethmcneil\_0', 'candys-outlet-store',  
'The Wires Zone', 'IQ Home Entertainment', 'Audiovideodirect',  
'BlueDealz', 'ditech\_store', 'cttax', 'theimagingworld',  
'Nationwide Distributor', 'DigitalShopper', 'great\_retailers',  
'daniellcreighto\_5', 'ami\_lamra', 'Shoplet', 'iban\_tech-8',  
'ovvapro', 'kourke', 'ezcellerz', 'serversourcetechnology',  
'autoplicity.com', 'Hotelectronics', 'jamhuritech', 'onecall',  
'salherman', 'business-supplies-2009', 'universe-deals',  
'livinitandlovinitretail', 'claimthis', 'aad\_77', 'gvinu',  
'pricebreak-deals', 'ProMax Commerce', 'balkan\_sale', 'saving2yo  
u',  
'MyOfficeInnovations', 'redcloudtech', 'SCONTROLC', 'ielectrica',  
'timespioneer2016', 'Simple Cell Inc', 'getdealonline', 'atoz123  
9',  
'Gigasonic', 'Quality Photo', 'dealsfellow',  
'Rocky Mountain Textbooks', 'gsalabama', 'TELeasy',  
'sunshinemeworld8689', 'turbopm', 'tanyastreasuries',  
'vminnovations', 'Digital Media Outlet', 'overstock5577',  
'eGoodies', 'volaudio', 'proavdealer', 'Concept Electronics',  
'globalavlcc', 'Spy Tec', 'department75', 'blueboxsales',  
'uShopMall', 'AngelSeller', '2011dailydeals', 'SaveTronix',  
'eternalphoto', 'mikromarc', 'ands3ts', 'extremeelectronix',  
'Overstock', 'pcRUSH', 'barterboxusa', 'gamegear14', 'Neobits',  
'zara4573', 'buy12volt', 'VIPOUTLET', 'iSave - Walmart.com',  
'dependix', 'Hippo-Deals', 'sound\_pros', 'summerishere-0',  
'TigerDirect - Walmart.com', 'tedselectronics1',  
'surplusstockpile', 'cyberoverstock-com', 'pillsboy3000',  
'dinsdale-usa', 'Amazing Cross Cloud Deals', 'XOTIC PC',  
'safeandsoundhq', 'California Electronics',  
'www-bestelectronicsoutlet-com', 'sandbid-com', 'thekeykey',  
'AAAA UNIVERSE', 'onlybargains2011', 'thrillingaudio',  
'CarAudioDistributors', 'Spddys LLC', 'ETM Online Deals, LLC',

'Aztek Computers', 'give5tocancer', 'Alwayz On Sale',  
'acumendiscs', 'Toolboxsupply', 'guru\_5', 'sqellc',  
'mainly-electronics', 'yogi-comp', 'discountparts99', 'bigdeals',  
'alohasgt', 'Zoro', 'Amazing Deals Online', 'Lastoftheloot',  
'Buy.com - Bestbuy.com', 'Hisense USA Corporation',  
'Computer Brain', 'Runtechmedia', 'kcla1026', '702-motoring',  
'Onlinecarstereo\_Store', 'happychicken-com', 'JBMO Sales',  
'e-shopsave', 'john-6807', 'digital-solomon', 'Always.Deals',  
'loothive', 'alex\_the\_terrible', 'secondipiety', 'PIXIBYTEs INC',  
'superdealunlimited', 'satman56', 'grliquidators16',  
'Rover Store, Inc.', 'Mobile Top Deal', 'Amazing Savings',  
'Wholesale Photo Supply Inc.', 'focuscamera', 'rabera',  
'cgelectronics16', 'pano\*rama', 'Elem Technologies',  
'anjcloseouts', 'hokie\_vt', 'onsale777', 'HiDEF Lifestyle',  
'amazingtexasdeals', 'Venu', 'tierlxswccx', 'Zoe Systems',  
'Good Luck Bargains', 'computeroverhauls', 'hollitronic',  
'irenzelio534', 'georgec4459', 'iSave', 'isellstuff86',  
'jakejake25', 'MegaWatts', 'Department75', 'prepre81',  
'Deals Unlimited', 'Digitaldjgear', 'hotdeals4u', 'itatlantic',  
'charlatan19', 'TEKENVY - Walmart.com', 'nitzafriedma\_0',  
'your-best-store', 'CopyThat!', 'radha\_stores', 'ltwright2006',  
'maxxal', 'northwest-outdoor', 'directnetshopdeal', 'DeepDiscoun  
t',  
'smart\_alliance', 'Widgetree', 'Moda Bella', 'discountsjungle',  
'Maxi-Aids', 'Product Source', 'jonrod\_b49gih', 'feeff',  
'hanabeauty10', 'book2301', 'PPG Consulting, Inc',  
'ZeeMo Distribution', 'shinjuku-donquixote', 'shopdivvy',  
'biction', 'Tunny LLC', 'MemoryC, Inc.', 'fanteller', 'cokekid88',  
'wirelessmoo', 'clocksntoys', 'pasbuy', 'qsccomputer',  
'umati\_housewtp', 'brake2buylc', 'Hot Deals Electronics',  
'sanhesun', 'Ritz Camera', 'trd\_digital', 'alpinespringers99',  
'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', 'omarafat09',  
'bidallies', 'Sabrent', 'ListenUp', 'Alto Music', 'accessoryhut2',  
'premium-electronics', 'videojam-half', 'hahiscool', 'Azano Inc',  
'aftermathoutdoor336', 'Electronics Plus', 'shipfast2day',  
'DiscountCentralOnline', 'techronics-store',  
'underground\_bargains', 'Advanced Gadget Inc.',  
'evolving-technologies', 'less.4.more', 'apc\_sales',  
'accessorynet', 'offers.store', 'Echo and Optics', 'jpkovacic',  
'elektech.1', 'jama2363', 'WAL-MART COM', '1topwire',  
'afdsolutions', 'us2014\_bacho', 'mobileshark',  
'United Liquidators', 'digjungle', 'vcc113',  
'Urban Armor Gear Official', 'dr.pepper357', 'DVCOM LLC',  
'jacquehuey817', 'Panda Rover', 'Roberts LP', 'mcmech',  
'Online King', 'bestelectronics1982', 'netroniques',  
'ProElectronics Distributing Inc.', 'kevin3sty', 'BrownBearTech',  
'TechMemoryStore', 'ritzcameras', 'newegg', 'altomusic',  
'Satellite Radio Superstore', 'daniedito', 'Firemall LLC',  
'big\_deals\_4\_you\_2', 'storesatisfactory', 'Top Rated Seller',  
'blucoil', 'SDC-Photo', 'Bluehaze', 'nzxtinc', 'abood.av',  
'buy.direct', 'josh-baskin-company', 'sobeonline1',  
'Beach Audio Inc - Walmart.com', 'torchehunte0', 'techmiaonline',  
'exclusivecharms', '428ypz', 'Mwave', 'tkservices',

```
'liquid.assets', 'TDPEP', 'frontrowelectronics',
'SSE Photo & Video', 'TopRatedSeller', 'VIP OUTLET', 'BulkBuys',
'moviegoddess85', 'tuminatr', 'ryanblue11', 'bensdealz7',
'ilikelike1', 'daylighttechnologies', 'nationwidedistributors',
'electr0nicsrus', 'nothingbutsavings', 'World Class Sales',
'Sports&Gadgets', 'einfinityshop4', 'Murloc Express',
'Savings Center', 'Buywireless'], dtype=object)
```

```
In [31]: dt_electronics.pricesmerchant.value_counts()
```

Out[31]:	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
	BuyVPC	21
	Car Audio Closeout	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
	Dell.com	10
	Newegg	10
	outletmp3	10
	Technology Galaxy	9
	Growkart	9
	Wholesale Connection	9
	Creative Audio	9
	Sound of Tri-State	9
	DemProductSales	9
	Clover Hill	9
	brandsmartusa	9
	wwstereo	9
	eLGeo Electronics	8
	HPP Enterprises	8
	My Goods	8
	gadget-town	8
	Hyper Microsystems, Inc.	8
	best_buy	7
	Tech For Less Inc	7
	Sam Ash Music Direct	7
	onecall	7
	extremenetwork	6
	DiamondBuy	6
	Made Technology	6
	Wayfair - Walmart.com	6
	Audiolab Stereo & Video Center	6
	wholesale_connection	6
	aimusa	5
	Datavision Computer Video	5
	ielectrica	5
	vipoutlet	5
	uShopMall	5
	sundown.one	5

Audiosavings	5
AV Toy Store	5
focuscamera	5
Walts TV	5
electronicsliquidators	5
Marine Discount Center, LLC	5
autosounds_warehouse	5
pcRUSH - Walmart.com	5
goldstar_tech	5
HappEshopper	5
tbdeals	5
www-sonicelectronix-com	5
refurbforless	5
ElectroCell	5
beachaudio	5
Monoprice Inc	4
Santa Monica Wireless	4
SaveTronix	4
TEKENVY - Walmart.com	4
a1Components	4
HookedOnTronics	4
Adorama Camera	4
Cameta Camera	4
evergreenpoint	4
Googol Shop	4
newopenbox	4
6Ave Express LLC	4
Hippo-Deals	4
SaveCentral, LLC	4
Hotelectronics	4
Mian Brothers Inc	4
HONESTDEALS	4
ProMax Commerce	4
redtagcamera	4
antonline	4
17th St Photo Supply Inc	4
autosounds4less	4
Computer Brain	3
n.n3288	3
ETM Online Deals, LLC	3
afdsolutions	3
6ave	3
BidDeal	3
Rover Store, Inc.	3
PT Etech	3
extremeelectronix	3
Always.Deals	3
Alex Magic Electronics	3
True Modern Electronics	3
Rocky Mountain Textbooks	3
poppy9780	3
Aztek Computers	3
Nexit, Inc.	3
GRAMOPHONE (Authorized Dealer)	3
dwgreatdeals	3
Audiovideodirect	3
tfldealz	3

OJCommerce	3
fanteller	3
gearbrokers	3
Wayfair	3
SCONTROLC	3
Circuit City - Walmart.com	3
DLR Hunter	3
VIPOUTLET	3
djwoodie	3
Adorama	3
Nationwide Distributor	3
hidedeflifestyle1	3
Tech2Date	3
videoandaudiocenter	3
listenup	3
Hisense USA Corporation	3
iphoneguy19	3
Simple Cell Inc	3
Pro AV Dealer	3
Samy's Camera	3
photovideo4less	3
CopyThat!	3
mygoods	2
jamhuritech	2
dwi-international-8	2
DigitalShopper	2
Global Audio Video LLC	2
Tri State Camera	2
aapltree	2
AUDIO WATT STORE	2
eCopUSA	2
Hot Deals 4 Less?	2
gpscity	2
gamegear14	2
MA Trade LLC	2
platinummicro	2
Music123	2
sando_distributions	2
Unbeatable Deals	2
book2301	2
OliveElect, LLC	2
lite_corp	2
pmiauctions	2
Kttrading14	2
The Wires Zone	2
raffaffle	2
Zoro	2
VMInnovations	2
superdealunlimited	2
District Camera	2
Olivabel	2
Mwave	2
Amazing Cross Cloud Deals	2
Deals Unlimited	2
EZ Tech	2
CarAudioDistributors	2
West Coast Limited	2

mnmdiscount1	2
sound_pros	2
ipadelectronics	2
timespioneer2016	2
thepixelhub	2
great_retailers	2
compucell.tech	2
DeepDiscount	2
TEKONLINE	2
TigerDirect - Walmart.com	2
GeekZilla	2
E-Warehouse-Deals	2
XOTIC PC	2
raghav_205	2
big_deals_4_you_2	2
DigitalKynect	2
digiherd	2
pcRUSH	2
pano*rama	2
altomusic	2
ericy23	2
JB Tool Sales	2
qsccomputer	2
Onlinecarstereo_Store	2
tri-state-camera	2
ravencat3	2
tunnyllc	2
Overstock	2
MemoryC, Inc.	2
Coldriver20 Electronics Inc	2
Always In Touch	2
universe-deals	2
ELAVA LLC	2
hokie_vt	2
Good Guys Electronics	2
Moda Bella	2
PPG Consulting, Inc	2
zara4573	2
guru_5	2
barterboxusa	2
volaudio	2
DiscountCentralOnline	2
Acumen Disc	2
ONE STOP SHOP	2
electr0nicsrus	2
globalavllc	2
iSave	2
atoz1239	2
serversourcetechnology	2
Digital Media Outlet	2
Shopemco	2
bigvalueinc	2
bling	2
Good Luck Bargains	2
iSave - Walmart.com	2
HiDEF Lifestyle	2
outletpc	2

dealonline18	2
Audio Advice	2
discountparts99	1
ryanblue11	1
Panda Rover	1
satman56	1
Electronics101deals	1
daylighttechnologies	1
mobileshark	1
blucoil	1
john-6807	1
accessoryhut2	1
gabrians	1
recommercedallas	1
einfinityshop4	1
shopcelldeals	1
amazingtexasdeals	1
yogi-comp	1
tkservices	1
cashinwhi	1
charlatan19	1
2011dailydeals	1
techronics-store	1
anjcloseouts	1
doitjust4fun	1
kays_place_too	1
singledadsunited	1
Alwayz On Sale	1
pillsboy3000	1
onlybargains2011	1
TopRatedSeller	1
accesscables	1
Xcess Limited	1
cybervalue	1
HiFiSoundConnection	1
e-shopsave	1
sunshineworld8689	1
nothingbutsavings	1
hypermicrosystems	1
autoplicity.com	1
less.4.more	1
vminnovations	1
Alto Music	1
getdealonline	1
1 Stop Electronics Center, Inc.	1
eGoodies	1
techmiaonline	1
nitzafriedma_0	1
Sabrent	1
Lastoftheloot	1
jbtoolsales02	1
Incredible Cellular LLC	1
livinitandlovininitretail	1
bluedealz-com	1
netroniques	1
Gigasonic	1
Real Basics	1

bensdealz7	1
Quality Photo	1
Music123 - Walmart.com	1
Mobile Top Deal	1
Triplenet Pricing	1
bornewinner	1
ezcellerz	1
bestelectronics1982	1
jama2363	1
Aiwa Inc.	1
cttax	1
Echo and Optics	1
secondipity	1
47st. Photo	1
summerishere-0	1
ovvapro	1
ami_lamra	1
wirelessmoo	1
JBMO Sales	1
Beach Audio Inc - Walmart.com	1
datu503	1
accessorynet	1
PIXIBYTES INC	1
amc0052010	1
zal-digital	1
itatlantic	1
new-technologies	1
gsalabama	1
Malelo and Company	1
World Class Sales	1
Hot Deals Electronics	1
safeandsoundhq	1
ritzcameras	1
sandbid-com	1
iTech INTL	1
istrawberrize	1
tedselectronics1	1
TDPEP	1
Buywireless	1
broomstick6	1
digjungle	1
proavdealer	1
aftermathoutdoor336	1
jakejake25	1
evolving-technologies	1
Give 5 To Cancer, Inc	1
Bluehaze	1
proaudio654	1
SharpPrices	1
Electronics Plus	1
305mia65	1
lwin	1
aad_77	1
Soltech US Corp	1
Zoe Systems	1
salherman	1
nationwidedistributors	1

brake2buylc	1
Yogi Comp	1
legendmicro-wholesale	1
saving2you	1
jpkovacic	1
pro_car_audio_consulting	1
ltwright2006	1
Azano Inc	1
Fab4Electronics	1
AAAA UNIVERSE	1
dieg5651_8xrznj2zi	1
Toolboxsupply	1
eternalphoto	1
MegaWatts	1
umpquablue	1
prepre81	1
Sports&Gadgets	1
happychicken-com	1
Digitaldjgear	1
d.a.n.t_tech	1
hotdeals4u	1
mainly-electronics	1
tanyastreasuries	1
zoneusa	1
biction	1
RackGo	1
luckeystore	1
onsale777	1
jbcuttlefish	1
ryajoh7	1
Wholesale Photo Supply Inc.	1
teampsales	1
bidallies	1
IQ Home Entertainment	1
bestserviceguys	1
thrillingaudio	1
department75	1
cokekid88	1
storesatisfactory	1
shipfast2day	1
liquid.assets	1
Amazing Savings	1
aSavings	1
shinjuku-donquixote	1
Buy.com - Bestbuy.com	1
stay_at_home_dads	1
Runtechmedia	1
alpinespringers99	1
blueboxsales	1
truemodernelectronicsusa	1
mpe007	1
elektech.1	1
robertscamera	1
digital-solomon	1
feeff	1
WeaKnees DVR Superstore	1
moviegoddess85	1

Deals Daily	1
apc_sales	1
georgec4459	1
TELeasy	1
theimagingworld	1
soundoftristate	1
mamastore_co	1
Savings Center	1
DVCOM LLC	1
weoutfit	1
turtlebeach_inc	1
grliquidators16	1
clocksntoys	1
LC Bargain Sales	1
closeoutmartdeals	1
K&M	1
business-supplies-2009	1
mikromarc	1
Neobits	1
maxxal	1
Top Rated Seller	1
OneStopShop	1
Widgetree	1
danielito	1
wirelessalliance	1
balkan_sale	1
umati_housewtp	1
gmatexpert	1
MyOfficeInnovations	1
Venu	1
thekeykey	1
WAL-MART COM	1
overstock5577	1
Murloc Express	1
Urban Armor Gear Official	1
sobeonline1	1
premium-electronics	1
dealsfellow	1
sanheshun	1
alex_the_terrible	1
gvinu	1
a4c	1
ilikeike1	1
kcla1026	1
ProElectronics Distributing Inc.	1
exclusivecharms	1
kourke	1
ELEM Technologies	1
onlycellphone	1
musicalsupplydirect	1
Spddys LLC	1
trd_digital	1
computeroverhauls	1
Direct Home Supplies	1
Top Deals	1
Concept Electronics	1
electric-flight-sales	1

dependix	1
pasbuy	1
us2014_bacho	1
vcc113	1
redcloudtech	1
torchehunte0	1
tuminatr	1
give5tocancer	1
alohasgt	1
California Electronics	1
irenzolio534	1
newegg	1
BuyNESP.com	1
Online King	1
webtron-x.com	1
ejuky	1
felixg7488	1
Roberts LP	1
hollitronic	1
topusedcomputers	1
buynmdeals	1
cynethmcneil_0	1
Product Source	1
buy.direct	1
hsum8885	1
surplusstockpile	1
lucky-31067	1
www-bestelectronicsoutlet-com	1
ditech_store	1
electricdiamond	1
directnetshopdeal	1
omarafat09	1
josh-baskin-company	1
Amazing Deals Online	1
Advanced Gadget Inc.	1
nzxtinc	1
discountsjungle	1
Firemall LLC	1
techhouse520	1
radha_stores	1
geekdeal	1
shopdivvy	1
All in Retail	1
jacquehuey817	1
TechMemoryStore	1
ands3ts	1
BulkBuys	1
hahiscool	1
nps	1
Department75	1
multitechinc	1
buy12volt	1
dr.pepper357	1
Platinum Micro	1
frontrowelectronics	1
cgelectronics16	1
candys-outlet-store	1

ListenUp	1
bargainbrokers	1
sqellc	1
BuilderDepot, Inc.	1
jazzman2275	1
Lightning Deals	1
ewholesale2011	1
mcmech	1
iban_tech-8	1
Maxi-Aids	1
MEMORY n MORE	1
Wayfair - Bestbuy.com	1
northwest-outdoor	1
offers.store	1
abood.av	1
CompuDirect	1
rabera	1
Shoplet	1
turbopm	1
hanabeauty10	1
Audio Visual Hub	1
tier1xswccx	1
daniellcreighto_5	1
SSE Photo & Video	1
smart_alliance	1
4surpluscity	1
videojam-half	1
g4mmlj	1
waltstv	1
ZeeMo Distribution	1
SDC-Photo	1
702-motoring	1
Spy Tec	1
Zeek and Zellies	1
tech1store	1
underground_bargains	1
jonrod_b49gih	1
19th Hole Deals	1
mike_gamesnmore	1
cable-n-chargers	1
dinsdale-usa	1
Tunny LLC	1
isellstuff86	1
race_choice	1
Satellite Radio Superstore	1
bigdeals	1
CosmeticMall	1
acumendiscs	1
rideausales	1
efesales	1
kevin3sty	1
chribrackma-0	1
axxii	1
1topwire	1
cyberoverstock-com	1
meiyaps	1
loothive	1

BuyDirect & Save	1
your-best-store	1
chrismron148	1
PhotoSavings	1
BrownBearTech	1
Double Deals	1
Circuit City	1
Focus Camera - Walmart.com	1
AngelSeller	1
VIP OUTLET	1
United Liquidators	1
Ritz Camera	1
428ypz	1
pricebreak-deals	1
BlueDealz	1
claimthis	1
playalegend	1
Name: pricesmerchant, dtype: int64	

```
In [32]: pricesmerchant_map=dt_electronics['pricesmerchant'].value_counts().to_dict()
pricesmerchant_map
```

```
Out[32]: {'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,
'BuyVPC': 21,
'Car Audio Closeout': 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,
'Dell.com': 10,
'Newegg': 10,
'outletmp3': 10,
'Technology Galaxy': 9,
'Growkart': 9,
'Wholesale Connection': 9,
'Creative Audio': 9,
'Sound of Tri-State': 9,
'DemProductSales': 9,
'Clover Hill': 9,
'brandsmartusa': 9,
'wwstereo': 9,
'eLGeo Electronics': 8,
'HPP Enterprises': 8,
'My Goods': 8,
'gadget-town': 8,
'Hyper Microsystems, Inc.': 8,
'best_buy': 7,
'Tech For Less Inc': 7,
'Sam Ash Music Direct': 7,
'onecall': 7,
'extremenetwork': 6,
'DiamondBuy': 6,
'Made Technology': 6,
'Wayfair - Walmart.com': 6,
'Audiolab Stereo & Video Center': 6,
'wholesale_connection': 6,
'aimusa': 5,
'Datavision Computer Video': 5,
'ielectrica': 5,
'vipoutlet': 5,
'uShopMall': 5,
'sundown.one': 5,
```

'Audiosavings': 5,  
'AV Toy Store': 5,  
'focuscamera': 5,  
'Walts TV': 5,  
'electronicsliquidators': 5,  
'Marine Discount Center, LLC': 5,  
'autosounds\_warehouse': 5,  
'pcRUSH - Walmart.com': 5,  
'goldstar\_tech': 5,  
'HappEshopper': 5,  
'tbdeals': 5,  
'www-sonicelectronix-com': 5,  
'refurbforless': 5,  
'ElectroCell': 5,  
'beachaudio': 5,  
'Monoprice Inc': 4,  
'Santa Monica Wireless': 4,  
'SaveTronix': 4,  
'TEKENVY - Walmart.com': 4,  
'a1Components': 4,  
'HookedOnTronics': 4,  
'Adorama Camera': 4,  
'Cameta Camera': 4,  
'evergreenpoint': 4,  
'Googol Shop': 4,  
'newopenbox': 4,  
'6Ave Express LLC': 4,  
'Hippo-Deals': 4,  
'SaveCentral, LLC': 4,  
'Hotelectronics': 4,  
'Mian Brothers Inc': 4,  
'HONESTDEALS': 4,  
'ProMax Commerce': 4,  
'redtagcamera': 4,  
'antonline': 4,  
'17th St Photo Supply Inc': 4,  
'autosounds4less': 4,  
'Computer Brain': 3,  
'n.n3288': 3,  
'ETM Online Deals, LLC': 3,  
'afdsolutions': 3,  
'6ave': 3,  
'BidDeal': 3,  
'Rover Store, Inc.': 3,  
'PT Etech': 3,  
'extremeelectronix': 3,  
'Always.Deals': 3,  
'Alex Magic Electronics': 3,  
'True Modern Electronics': 3,  
'Rocky Mountain Textbooks': 3,  
'poppy9780': 3,  
'Aztek Computers': 3,  
'Nexit, Inc.': 3,  
'GRAMOPHONE (Authorized Dealer)': 3,  
'dwgreatdeals': 3,  
'Audiovideodirect': 3,  
'tfldealz': 3,

'OJCommerce': 3,  
'fanteller': 3,  
'gearbrokers': 3,  
'Wayfair': 3,  
'SCONTROLC': 3,  
'Circuit City - Walmart.com': 3,  
'DLR Hunter': 3,  
'VIPOUTLET': 3,  
'djwoodie': 3,  
'Adorama': 3,  
'Nationwide Distributor': 3,  
'hideflifestyle1': 3,  
'Tech2Date': 3,  
'videoandaudiocenter': 3,  
'listenup': 3,  
'Hisense USA Corporation': 3,  
'iphoneguy19': 3,  
'Simple Cell Inc': 3,  
'Pro AV Dealer': 3,  
"Samy's Camera": 3,  
'photovideo4less': 3,  
'CopyThat!': 3,  
'mygoods': 2,  
'jamhuritech': 2,  
'dwi-international-8': 2,  
'DigitalShopper': 2,  
'Global Audio Video LLC': 2,  
'Tri State Camera': 2,  
'aapltree': 2,  
'AUDIO WATT STORE': 2,  
'eCopUSA': 2,  
'Hot Deals 4 Less?': 2,  
'gpscity': 2,  
'gamegear14': 2,  
'MA Trade LLC': 2,  
'platinummicro': 2,  
'Music123': 2,  
'sando\_distributions': 2,  
'Unbeatible Deals': 2,  
'book2301': 2,  
'OliveElect, LLC': 2,  
'lite\_corp': 2,  
'pmiauctions': 2,  
'kttrading14': 2,  
'The Wires Zone': 2,  
'raffaffle': 2,  
'Zoro': 2,  
'VMInnovations': 2,  
'superdealunlimited': 2,  
'District Camera': 2,  
'Olivabel': 2,  
'Mwave': 2,  
'Amazing Cross Cloud Deals': 2,  
'Deals Unlimited': 2,  
'EZ Tech': 2,  
'CarAudioDistributors': 2,  
'West Coast Limited': 2,

'mnmdiscount1': 2,  
'sound\_pros': 2,  
'ipadelectronics': 2,  
'timespioneer2016': 2,  
'thepixelhub': 2,  
'great\_retailers': 2,  
'compucell.tech': 2,  
'DeepDiscount': 2,  
'TEKONLINE': 2,  
'TigerDirect - Walmart.com': 2,  
'GeekZilla': 2,  
'E-Warehouse-Deals': 2,  
'XOTIC PC': 2,  
'raghav\_205': 2,  
'big\_deals\_4\_you\_2': 2,  
'DigitalKynect': 2,  
'digiherd': 2,  
'pcRUSH': 2,  
'pano\*rama': 2,  
'altomusic': 2,  
'ericy23': 2,  
'JB Tool Sales': 2,  
'qsccomputer': 2,  
'Onlinecarstereo\_Store': 2,  
'tri-state-camera': 2,  
'ravencat3': 2,  
'tunnyllc': 2,  
'Overstock': 2,  
'MemoryC, Inc.': 2,  
'Coldriver20 Electronics Inc': 2,  
'Always In Touch': 2,  
'universe-deals': 2,  
'ELAVA LLC': 2,  
'hokie\_vt': 2,  
'Good Guys Electronics': 2,  
'Moda Bella': 2,  
'PPG Consulting, Inc': 2,  
'zara4573': 2,  
'guru\_5': 2,  
'barterboxusa': 2,  
'volaudio': 2,  
'DiscountCentralOnline': 2,  
'Acumen Disc': 2,  
'ONE STOP SHOP': 2,  
'electr0nicsrus': 2,  
'globallavllc': 2,  
'iSave': 2,  
'atoz1239': 2,  
'serversourcetechnology': 2,  
'Digital Media Outlet': 2,  
'Shopemco': 2,  
'bigvalueinc': 2,  
'bling': 2,  
'Good Luck Bargains': 2,  
'iSave - Walmart.com': 2,  
'HiDEF Lifestyle': 2,  
'outletpc': 2,

'dealonline18': 2,  
'Audio Advice': 2,  
'discountparts99': 1,  
'ryanblue11': 1,  
'Panda Rover': 1,  
'satman56': 1,  
'Electronics101deals': 1,  
'daylighttechnologies': 1,  
'mobileshark': 1,  
'blucoil': 1,  
'john-6807': 1,  
'accessoryhut2': 1,  
'gabrians': 1,  
'recommercedallas': 1,  
'einfinityshop4': 1,  
'shopcelldeals': 1,  
'amazingtexasdeals': 1,  
'yogi-comp': 1,  
'tkservices': 1,  
'cashinwhi': 1,  
'charlatan19': 1,  
'2011dailydeals': 1,  
'techronics-store': 1,  
'anjcloseouts': 1,  
'doitjust4fun': 1,  
'kays\_place\_too': 1,  
'singledadsunited': 1,  
'Alwayz On Sale': 1,  
'pillsboy3000': 1,  
'onlybargains2011': 1,  
'TopRatedSeller': 1,  
'accesscables': 1,  
'Xcess Limited': 1,  
'cybervalue': 1,  
'HiFiSoundConnection': 1,  
'e-shopsave': 1,  
'sunshineworld8689': 1,  
'nothingbut savings': 1,  
'hypermicrosystems': 1,  
'autopllicity.com': 1,  
'less.4.more': 1,  
'vminnovations': 1,  
'Alto Music': 1,  
'getdealonline': 1,  
'1 Stop Electronics Center, Inc.': 1,  
'eGoodies': 1,  
'techmiaonline': 1,  
'nitzafriedma\_0': 1,  
'Sabrent': 1,  
'Lastoftheloot': 1,  
'jbtoolsales02': 1,  
'Incredible Cellular LLC': 1,  
'livinitandlovinitretail': 1,  
'bluedealz-com': 1,  
'netroniques': 1,  
'Gigasonic': 1,  
'Real Basics': 1,

'bensdealz7': 1,  
'Quality Photo': 1,  
'Music123 - Walmart.com': 1,  
'Mobile Top Deal': 1,  
'Triplenet Pricing': 1,  
'bornewinner': 1,  
'ezcellerz': 1,  
'bestelectronics1982': 1,  
'jama2363': 1,  
'Aiwa Inc.': 1,  
'cttax': 1,  
'Echo and Optics': 1,  
'secondipity': 1,  
'47st. Photo': 1,  
'summerishere-0': 1,  
'ovvapro': 1,  
'ami\_lamra': 1,  
'wirelessmoo': 1,  
'JBMO Sales': 1,  
'Beach Audio Inc - Walmart.com': 1,  
'datu503': 1,  
'accessorynet': 1,  
'PIXIBYTES INC': 1,  
'amc0052010': 1,  
'zal-digital': 1,  
'itatlantic': 1,  
'new-technologies': 1,  
'gsalabama': 1,  
'Malelo and Company': 1,  
'World Class Sales': 1,  
'Hot Deals Electronics': 1,  
'safeandsoundhq': 1,  
'ritzcameras': 1,  
'sandbid-com': 1,  
'iTech INTL': 1,  
'istrawberrize': 1,  
'tedselectronics1': 1,  
'TDPEP': 1,  
'Buywireless': 1,  
'broomstick6': 1,  
'digjungle': 1,  
'proavdealer': 1,  
'aftermathoutdoor336': 1,  
'jakejake25': 1,  
'evolving-technologies': 1,  
'Give 5 To Cancer, Inc': 1,  
'Bluehaze': 1,  
'proaudio654': 1,  
'SharpPrices': 1,  
'Electronics Plus': 1,  
'305mia65': 1,  
'1win': 1,  
'aad\_77': 1,  
'Soltech US Corp': 1,  
'Zoe Systems': 1,  
'salherman': 1,  
'nationwidedistributors': 1,

'brake2buylc': 1,  
'Yogi Comp': 1,  
'legendmicro-wholesale': 1,  
'saving2you': 1,  
'jpkovacic': 1,  
'pro\_car\_audio\_consulting': 1,  
'ltwright2006': 1,  
'Azano Inc': 1,  
'Fab4Electronics': 1,  
'AAAAA UNIVERSE': 1,  
'dieg5651\_8xrznj2zi': 1,  
'Toolboxsupply': 1,  
'eternalphoto': 1,  
'MegaWatts': 1,  
'umpquablue': 1,  
'prepre81': 1,  
'Sports&Gadgets': 1,  
'happychicken-com': 1,  
'Digitaldjgear': 1,  
'd.a.n.t\_tech': 1,  
'hotdeals4u': 1,  
'mainly-electronics': 1,  
'tanyastreasuries': 1,  
'zoneusa': 1,  
'biction': 1,  
'RackGo': 1,  
'luckeyystore': 1,  
'onsale777': 1,  
'jbcuttlefish': 1,  
'ryajoh7': 1,  
'Wholesale Photo Supply Inc.': 1,  
'teampsales': 1,  
'bidallies': 1,  
'IQ Home Entertainment': 1,  
'bestserviceguys': 1,  
'thrillingaudio': 1,  
'department75': 1,  
'cokekid88': 1,  
'storesatisfactory': 1,  
'shipfast2day': 1,  
'liquid.assets': 1,  
'Amazing Savings': 1,  
'aSavings': 1,  
'shinjuku-donquixote': 1,  
'Buy.com - Bestbuy.com': 1,  
'stay\_at\_home\_dads': 1,  
'Runtechmedia': 1,  
'alpinespringers99': 1,  
'blueboxsales': 1,  
'truemodernelectronicsusa': 1,  
'mpe007': 1,  
'elektech.1': 1,  
'robertscamera': 1,  
'digital-solomon': 1,  
'feefff': 1,  
'WeaKnees DVR Superstore': 1,  
'moviegoddess85': 1,

'Deals Daily': 1,  
'apc\_sales': 1,  
'georgec4459': 1,  
'TELeasy': 1,  
'theimagingworld': 1,  
'soundoftristate': 1,  
'mamastore\_co': 1,  
'Savings Center': 1,  
'DVCOM LLC': 1,  
'weoutfit': 1,  
'turtlebeach\_inc': 1,  
'grliquidators16': 1,  
'clocksntoys': 1,  
'LC Bargain Sales': 1,  
'closeoutmartdeals': 1,  
'K&M': 1,  
'business-supplies-2009': 1,  
'mikromarc': 1,  
'Neobits': 1,  
'maxxa1': 1,  
'Top Rated Seller': 1,  
'OneStopShop': 1,  
'Widgetree': 1,  
'daniedito': 1,  
'wirelessalliance': 1,  
'balkan\_sale': 1,  
'umati\_housewtp': 1,  
'gmatexpert': 1,  
'MyOfficeInnovations': 1,  
'Venu': 1,  
'thekeykey': 1,  
'WAL-MART COM': 1,  
'overstock5577': 1,  
'Murloc Express': 1,  
'Urban Armor Gear Official': 1,  
'sobeonline1': 1,  
'premium-electronics': 1,  
'dealsfellow': 1,  
'sanheshun': 1,  
'alex\_the\_terrible': 1,  
'gvinu': 1,  
'a4c': 1,  
'ilikeikel': 1,  
'kcla1026': 1,  
'ProElectronics Distributing Inc.': 1,  
'exclusivecharms': 1,  
'kourke': 1,  
'Elem Technologies': 1,  
'onlycellphone': 1,  
'musicalsupplydirect': 1,  
'Spddys LLC': 1,  
'trd\_digital': 1,  
'computeroverhauls': 1,  
'Direct Home Supplies': 1,  
'Top Deals': 1,  
'Concept Electronics': 1,  
'electric-flight-sales': 1,

'dependency': 1,  
'pasbuy': 1,  
'us2014\_bacho': 1,  
'vcc113': 1,  
'redcloudtech': 1,  
'torchehunte0': 1,  
'tuminatr': 1,  
'give5tocancer': 1,  
'alohasgt': 1,  
'California Electronics': 1,  
'irenzolio534': 1,  
'newegg': 1,  
'BuyNESP.com': 1,  
'Online King': 1,  
'webtron-x.com': 1,  
'ejuky': 1,  
'felixg7488': 1,  
'Roberts LP': 1,  
'hollitronic': 1,  
'topusedcomputers': 1,  
'buynmdeals': 1,  
'cynethmcneil\_0': 1,  
'Product Source': 1,  
'buy.direct': 1,  
'hsum8885': 1,  
'surplusstockpile': 1,  
'lucky-31067': 1,  
'www-bestelectronicsoutlet-com': 1,  
'ditech\_store': 1,  
'electricdiamond': 1,  
'directnetshopdeal': 1,  
'omarafat09': 1,  
'josh-baskin-company': 1,  
'Amazing Deals Online': 1,  
'Advanced Gadget Inc.': 1,  
'nzxtinc': 1,  
'discountsjungle': 1,  
'Firemall LLC': 1,  
'techhouse520': 1,  
'radha\_stores': 1,  
'geekdeal': 1,  
'shopdivvy': 1,  
'All in Retail': 1,  
'jacquehuey817': 1,  
'TechMemoryStore': 1,  
'ands3ts': 1,  
'BulkBuys': 1,  
'hahiscool': 1,  
'nps': 1,  
'Department75': 1,  
'multitechinc': 1,  
'buy12volt': 1,  
'dr.pepper357': 1,  
'Platinum Micro': 1,  
'frontrowelectronics': 1,  
'cgelectronics16': 1,  
'candys-outlet-store': 1,

'ListenUp': 1,  
'bargainbrokers': 1,  
'sqellc': 1,  
'BuilderDepot, Inc.': 1,  
'jazzman2275': 1,  
'Lightning Deals': 1,  
'ewholesale2011': 1,  
'mcmech': 1,  
'iban\_tech-8': 1,  
'Maxi-Aids': 1,  
'MEMORY n MORE': 1,  
'Wayfair - Bestbuy.com': 1,  
'northwest-outdoor': 1,  
'offers.store': 1,  
'abood.av': 1,  
'CompuDirect': 1,  
'rabera': 1,  
'Shoplet': 1,  
'turbopm': 1,  
'hanabeauty10': 1,  
'Audio Visual Hub': 1,  
'tierlxswccx': 1,  
'daniellcreighto\_5': 1,  
'SSE Photo & Video': 1,  
'smart\_alliance': 1,  
'4surpluscity': 1,  
'videojam-half': 1,  
'g4mmlj': 1,  
'waltstv': 1,  
'ZeeMo Distribution': 1,  
'SDC-Photo': 1,  
'702-motoring': 1,  
'Spy Tec': 1,  
'Zeek and Zellies': 1,  
'tech1store': 1,  
'underground\_bargains': 1,  
'jonrod\_b49gih': 1,  
'19th Hole Deals': 1,  
'mike\_gamesnmore': 1,  
'cable-n-chargers': 1,  
'dinsdale-usa': 1,  
'Tunny LLC': 1,  
'isellstuff86': 1,  
'race\_choice': 1,  
'Satellite Radio Superstore': 1,  
'bigdeals': 1,  
'CosmeticMall': 1,  
'acumendiscs': 1,  
'rideausales': 1,  
'efesales': 1,  
'kevin3sty': 1,  
'chribrackma-0': 1,  
'axxii': 1,  
'1topwire': 1,  
'cyberoverstock-com': 1,  
'meiyaps': 1,  
'loothive': 1,

```
'BuyDirect & Save': 1,  
'your-best-store': 1,  
'chrismadron148': 1,  
'PhotoSavings': 1,  
'BrownBearTech': 1,  
'Double Deals': 1,  
'Circuit City': 1,  
'Focus Camera - Walmart.com': 1,  
'AngelSeller': 1,  
'VIP OUTLET': 1,  
'United Liquidators': 1,  
'Ritz Camera': 1,  
'428ypz': 1,  
'pricebreak-deals': 1,  
'BlueDealz': 1,  
'claimthis': 1,  
'playalegend': 1}
```

here I will replace 'Bestbuy.com' with 2109

```
In [33]: dt_electronics['pricesmerchant']=dt_electronics['pricesmerchant'].map(pricesmerchant_map)
dt_electronics.head()
```

Out[33]:

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

This column indicate shipping charges of each delivery

```
In [34]: dt_electronics.pricesshipping.unique()
```

```
Out[34]: 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', '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',
   'Free Shipping for this Item', 'USD 5.00', 'USD 140.00',
   'USD 7.25', 'USD 14.00', 'USD 35.03', 'USD 0.99', 'USD 25.76',
   'USD 6.00', 'Free Next Day Delivery (USA)', 'USD 2.99',
   'USD 20.00', 'USD 22.13', 'USD 60.00', 'USD 2.85', 'USD 5.99'],
  dtype=object)
```

I will replace those values which indicate free delivery with 0

```
In [35]: dt_electronics['pricesshipping']=dt_electronics['pricesshipping'].replace
(['Free Expedited Shipping for most orders over $49','Free Shipping on orders 35 and up','Free Expedited Shipping','Free Delivery','Free Shipping on 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 [36]: dt_electronics.pricesshipping.unique()
```

```
Out[36]: array([nan, 0, 'Freight', 'Standard', 'Value', 'Expedited', 'USD 3.99',
   'USD 13.32', 'Shipping Charges Apply', '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], dtype=object)
```

Now I will remove the USD string from each value

```
In [37]: 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  
    ,16.00,12.78,13.60,11.30,6.65,15.42,18.61,4.88,10.00,13.61,25.00,150.00,1  
    96.92,10.10,7.81,24.04,22.64,75.00,16.45,179.35,27.94,15.68,37.65,18.60,1  
    3.95,79.42,5.00,140.00,7.25,14.00,35.03,0.99,25.76,6.00,2.99,20.00,22.13,  
    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 [38]: dt_electronics['pricesshipping']=dt_electronics['pricesshipping'].replace(['Freight','Standard','Value','Expedited','Shipping Charges Apply'],np.nan)
```

```
In [39]: dt_electronics.pricesshipping.unique()
```

```
Out[39]: array([  nan,   0.  ,   3.99,  13.32,   3.39,  16.  ,  12.78,  13.6 ,  
    11.3 ,   6.65,  15.42,  18.61,   4.88,  10.  ,  13.61,  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,  79.42,   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 [40]: mean=dt_electronics.pricesshipping.mean()  
mean=round(mean)
```

```
In [41]: dt_electronics[['pricesshipping']] = dt_electronics[['pricesshipping']].fillna(mean, axis=1)
```

```
In [42]: dt_electronics['pricesshipping']=dt_electronics['pricesshipping'].astype('int64')
```

```
In [43]: dt_electronics.priceshipping.unique()
```

```
Out[43]: 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 [44]: `dt_electronics.brand.unique()`

```
Out[44]: array(['Grace Digital', 'Lenovo', 'House of Marley', 'Sony', 'Acer', 'LG',
   'Yamaha', 'Kicker', 'Turtle Beach', 'Russound', 'Sandisk',
   'Logitech', 'Samsung', 'WD', 'Nikon', 'ASUS', 'Thermaltake',
   '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', 'Zoo
   m',
   'Sennheiser', 'DENAQ', 'Sanus', 'Bose®', 'Western Digital',
   'Sling Media', 'Sima', 'Netgear', 'V-MODA', 'Manfrotto', 'Monste
   r',
   'Zmodo', 'VIZIO', 'Boytone', 'Siriusxm', 'ECOXGEAR', 'MartinLoga
   n',
   'XFX', 'HP', 'Olympus', 'Microsoft', 'Corsair', 'Wacom', 'Dual',
   'ZTE', 'V-Moda', 'Pyle Pro', 'Cerwin-Vega', 'Onkyo', 'NVIDIA',
   'CLARITY-TELECOM', 'Asus', 'iSimple', 'Definitive Technology',
   'Linksys', 'PyleHome', 'Kenwood', 'Braven', 'Jbl', 'MSI',
   'Lowepro', 'G-Technology', 'Optoma', 'Toshiba', 'GEKO',
   'Sideclick', 'Actiontec', 'ViewSonic', 'Buffalo',
   'Cobra Electronics', 'Siliconust', 'SunBriteTV', 'TP-Link',
   'Cooler Master', 'VINPOWER DIGITAL DIRECTSHIP', 'Midland',
   'Visidec', 'Bower', 'Sabrent', 'Klipsch Xr8i In-Ear Headphones',
   'Jabra', 'mophie', 'BIC America', 'Peak Design', 'Jensen',
   'Marantz', 'Belkin Inc.', 'Viper', 'AudioQuest', 'Case Logic',
   'Sound Design', 'kate spade new york', 'Buffalo Technology',
   'Bose', 'GoPro', 'KICKER', 'Tivo', 'Insignia', 'RCA', 'Beats',
   'Pyle', 'Tiffen', 'Secur', 'Sol Republic', '360fly', 'Steelserie
   s',
   'Mevo', 'Aiwa', 'Sherwood', 'MEE audio', 'Peerless-AV', 'Ghostek',
   'Epson', 'Sigma', 'ZAGG', 'Kingston', 'MTX Audio', 'Spartan',
   'Speck', 'Atrend', 'TCL', 'Sengled', 'Memorex', 'Antec',
   'Rand McNally', 'StarTech', 'Aftershokz', 'MTX', 'Novatel',
   'Google', 'LifeProof', 'SANUS', 'Innovative Technology', 'AOC',
   'Arris', 'PNY', 'Master Dynamic', 'Aluratek', 'PELICAN',
   'Digipower', 'House Of Marley', 'Anker', 'TiVo', 'VisionTek',
   'BOYO', 'Transcend', 'Jaybird', 'Urban Armor Gear', 'NZXT',
   'IOGEAR', 'Audioquest', 'SpeakerCraft', 'SKB', 'I.am+',
   'Planet Audio', 'FUGOO', 'Zubie', 'SHARKK', 'Goal Zero',
   'PreSonus', 'Mee Audio', 'Niles', 'Fusion Marine', 'NETGEAR',
   'M-Audio', 'Sunpak', 'ONLINE', 'Fitbit', 'Papago', 'Sonos', 'Evg
   a',
   'Tenba', 'X6D', 'ECO STYLE', 'Outdoor Tech', 'Cobra',
   'DENON - HEOS', 'CybertronPC', 'Chief', 'Ultimate Ears', 'JVC',
   'Diamond Multimedia', 'Hoya', 'PANAMX', 'Intel', 'D-Link',
   'Sdi Technologies, Inc.', 'BenQ', 'Master & Dynamic', 'Lorex',
   'Crosley', 'KEF', 'Thule', 'Motorola', 'Sonax', 'Leef',
   'Hauppauge', 'AfterShokz', 'Incipio', 'OmniMount', 'Victrola',
   'PYLE', 'Insignia™', 'Skullcandy', 'Skybell', 'Energizer',
   'j5create', 'VXi', 'Retrak', 'DreamWave'], dtype=object)
```

In [45]: `dt_electronics.brand.value_counts()`

Out[45]:	Sony	578
	Samsung	565
	Apple	188
	Yamaha	183
	Pioneer	137
	LG	132
	Logitech	82
	Lenovo	80
	Kenwood	74
	SanDisk	74
	CORSAIR	73
	WD	72
	Alpine	72
	Sennheiser	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
	Canon	34
	Olympus	34
	Definitive Technology	33
	Kicker	32
	V-Moda	32
	Epson	31
	Microsoft	28
	Grace Digital	27
	ECOXGEAR	25
	AudioQuest	25
	Actiontec	24
	DENAQ	24
	SOL REPUBLIC	24
	Sharp	23
	Joby	23
	Siriusxm	22
	Asus	22
	Kanto Living	21
	Zoom	21
	MEE audio	21
	Acer	21
	Nikon	20
	AOC	20

Jbl	20
Sylvania	19
Hisense	19
Klipsch	19
MartinLogan	18
Pyle Pro	18
Sanus	18
ViewSonic	18
V-MODA	18
360fly	18
Steelseries	17
Cooler Master	17
Insignia	16
Bose	16
Optoma	16
MSI	16
iSimple	16
Boytone	15
GEKO	15
Toshiba	15
Bose®	15
Pny	15
Thermaltake	15
Peak Design	15
Rand McNally	14
Bower	14
Polk Audio	14
NZXT	14
mophie	14
Linksys	13
HP	13
Isimple	13
Silicondust	13
StarTech.com	13
Manfrotto	13
Kingston	13
Belkin	13
Papago	13
SunBriteTV	13
Kanto	13
SVS	12
VIZIO	12
Wacom	12
Belkin Inc.	12
Marantz	12
Pyle	12
Westone	11
BIC America	11
NETGEAR	11
LifeProof	11
Jensen	11
Zmodo	11
G-Technology	11
Jvcm	10
Spartan	10
Sima	10
Visidec	10

RCA	10
Panamax	10
Google	10
Atrend	10
ZTE	10
MTX	10
Cobra Electronics	10
Sigma	9
RODE	9
Intel	9
Monster	9
Braven	9
SiriusXM	9
Sherwood	9
XFX	9
PyleHome	9
Mevo	9
Peerless-AV	9
Jaybird	8
SHARKK	8
Sound Design	8
Tiffen	8
Motorola	8
Midland	8
Escort	8
Arris	8
Tenba	8
Sengled	8
Niles	8
Dual	8
ZAGG	8
DENON - HEOS	7
Peerless	7
Motorola Home	7
CybertronPC	7
Viper	7
PYLE	7
GoPro	7
Cerwin Vega	7
VisionTek	7
House Of Marley	7
NVIDIA	6
Cobra	6
Polaroid	6
TCL	6
Sideclick	6
Aluratek	6
Speck	6
StarTech	6
PNY	6
PELICAN	6
Anker	6
SKB	6
CLARITY-TELECOM	6
Cerwin-Vega	6
SANUS	5
Tivo	5

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

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

I will get this rank with respect to price column

```
In [46]: ordinal_labels=dt_electronics.groupby(['brand'])['price'].mean().sort_values().index
```

```
In [47]: ordinal_labels
```

```
Out[47]: Index(['Retrak', 'Sunpak', 'Secur', 'iLive', 'Energizer', 'RCA', 'iSimple',
       'Belkin Inc.', 'DENAQ', 'Case Logic',
       ...,
       'Apple', 'Samsung', 'Optoma', 'Alienware', 'Asus', 'Marantz', 'MS
I',
       'Intel', 'LG', 'SunBriteTV'],
      dtype='object', name='brand', length=254)
```

```
In [48]: enumerate(ordinal_labels,0)
```

```
Out[48]: <enumerate at 0x2025bbffe58>
```

```
In [49]: ordinal_labels2={k:i for i,k in enumerate(ordinal_labels,0)}  
ordinal_labels2
```

```
Out[49]: {'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,  
'Jvcm': 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,  
'Pny': 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,  
'Master & Dynamic': 230,  
'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 [50]: dt_electronics['brand_ordinal_labels']=dt_electronics['brand'].map(ordinal_labels2)
dt_electronics.head()
```

Out[50]:

	pricesavailability	pricescondition	pricesdateSeen	pricesmerchant	pricesshi
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	2017-09-09T05:00:00Z,2017-18T13:00:00Z,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 [51]: dt_electronics['dateUpdated'] = pd.to_datetime(dt_electronics['dateUpdate
d'].astype('datetime64[ns]'),format='%y%m%d')
```

```
In [52]: dt_electronics['dateAdded'] =pd.to_datetime(dt_electronics['dateAdded'].a
stype('datetime64[ns]'),format='%y%m%d')
```

Now i create month and year column from this date field

```
In [53]: dt_electronics['Month_dtadd']=pd.DatetimeIndex(dt_electronics['dateAdded']).month  
dt_electronics['Year_dtadd']=pd.DatetimeIndex(dt_electronics['dateAdded']).year
```

```
In [54]: dt_electronics['Month_updated']=pd.DatetimeIndex(dt_electronics['dateUpdated']).month  
dt_electronics['Year_updated']=pd.DatetimeIndex(dt_electronics['dateUpdated']).year
```

```
In [55]: dt_electronics.head()
```

Out[55]:

	pricesavailability	pricescondition	pricesdateSeen	pricesmerchant	pricesshi
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	

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

```
In [56]: dt_electronics.drop(['dateAdded', 'dateUpdated'], axis=1, inplace=True)
```

```
In [57]: dt_electronics.head()
```

Out[57]:

	pricesavailability	pricescondition	pricesdateSeen	pricesmerchant	pricesshi
0	1590	4645	2017-05-05T10T20:00:00Z,2017-09T15:00:00Z	2109	
1	1590	4645	2017-10-08T10T02:00:00Z,2017-12T03:00:00Z,2017...	2109	
2	1590	4645	2017-10-09T10T19:00:00Z,2017-12T14:00:00Z,2017...	2109	
3	1590	4645	2017-09-09T08T05:00:00Z,2017-18T13:00:00Z,2017...	2109	
4	63	4645	2017-12-05T13:00:00Z	1119	

```
In [58]: dt_electronics.manufacturer.unique()
```

```
Out[58]: 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', 'Elitescreens',  
   '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', 'KLIPSCH', '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', 'Sunpa  
k',  
   'Spacepole Ince', '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', 'Pyl  
e',  
   'Alpine Electronics USA', 'Razer USA', 'niles', 'Lorex',  
   'Crosley Radio', 'Thule', 'MTRLC LLC',  
   'Jvc Real Sound System Z Series Ha-Sz2000 Stream Woofer Db Headpho  
nes (black, Japan Import) V-Moda',  
   'Klipsch', 'Sonax', 'eTape16', 'Motorola', 'Dantona', 'OmniMount',  
   'LowePro', 'Niles', 'Lenovo', 'Energizer', 'Emerge'], dtype=objc  
t)
```

```
In [59]: dt_electronics.manufacturer.value_counts()
```

Out[59]:

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	
Garmin	
24	
Parts	
24	
Western Digital	
24	
YAMAHA	
23	
Canon	
23	
Logitech	
23	
Samsung Electronics	
22	
Dell	
21	
NETGEAR	
21	
Zoom	
21	
Panasonic	

20  
Denaq  
20  
ONKYO  
20  
Acer  
19  
House of Marley  
18  
Power Acoustik  
18  
Manufacturers Limited Warranty:2 Year  
18  
SiriusXM  
18  
120  
18  
Yamaha Electronics  
17  
Joby  
16  
Kanto  
15  
Manufacturers Limited Warranty:3 Year  
15  
iSimple  
15  
Samsung IT  
15  
Grace Digital  
15  
Asus  
14  
Sylvania  
14  
Sanus  
13  
StarTech.com  
13  
Manufacturers Limited Warranty:Lifetime  
13  
SiliconDust USA, Inc  
13  
Bose  
13  
SanDisk  
13  
Razer  
13  
Papago  
13  
MTX  
13  
Manfrotto  
13  
Grace Digital Audio  
13

Logitech, Inc  
12  
Lenovo Corporate  
12  
Belkin  
12  
Vizio  
12  
JVC  
12  
Kenwood  
12  
Audioquest  
11  
TP-Link  
11  
Manufacturers Limited Warranty:1 Year  
11  
Boytone  
11  
OtterBox  
11  
Ecoxgear  
11  
BIC America  
11  
iHome  
11  
RAZER  
11  
Jensen  
11  
Westone Laboratories, Inc.  
11  
NZXT  
11  
Sony Mobile Communications, (USA) Inc  
10  
RCA  
10  
House Of Marley  
10  
SAMSUNG  
10  
LG ODD  
10  
Sima  
10  
ISIMPLE  
10  
KICKER  
10  
Audiovox  
9  
Canon Cameras US  
9  
Netgear Inc

9  
RODE  
9  
Sherwood  
9  
JAYBIRD  
8  
Dual  
8  
Zagg  
8  
1 year warranty  
8  
Jbl Everest 110 In-Ear Wireless Bluetooth Headphones (silver)  
8  
Tiffen  
8  
AudioQuest  
8  
Midland  
8  
Sharp  
8  
Sigma  
7  
Cybertronpc  
7  
JBL  
7  
Pyle  
7  
Spacepole Ince  
7  
MEE audio  
7  
Peerless  
7  
PYLE  
7  
Apple Computer  
7  
Visiontek  
7  
Viper by GLD Products  
7  
Microsoft  
7  
Peak Design  
7  
Lenovo Idea  
7  
5 Years  
6  
ELITE SCREENS DIRECTSHIP  
6  
LowePro  
6

Speck  
6  
Elite Screens  
6  
ARRIS Solutions, Inc.  
6  
SKB Cases  
6  
1 Year  
6  
Allround Software  
6  
Apple Computer (Direct)  
6  
Klipsch  
6  
Coolermaster  
6  
Zmodo  
6  
MTRLC LLC  
6  
BOSE  
6  
Pelican  
6  
Aluratek  
6  
Spartan  
6  
Limited Lifetime Warranty  
6  
Polaroid  
6  
KLIPSCH  
6  
Bower Camera  
5  
eTape16  
5  
Marantz  
5  
Buffalo Technology  
5  
Incipio  
5  
Sonos  
5  
Novatel  
5  
Ilive Blue  
5  
Boyo  
5  
TiVo  
5  
Urban Armor Gear

5  
Sound Around  
5  
Cerwin Vega  
4  
Jvc Real Sound System Z Series Ha-Sz2000 Stream Woofer Db Headphones (bla  
ck, Japan Import) V-Moda 4  
Pac  
4  
Klipsch X12 Bluetooth Neckband Headphones (black Leather)  
4  
Fusion  
4  
Siriusxm  
4  
StarTech  
4  
Bower  
4  
Emerge  
4  
2 Years  
4  
EPSON AMERICA, INC.  
4  
Sonic Alert  
4  
Thule  
4  
Dantona  
4  
TRANSCEND Storage  
4  
Mophie  
4  
Go-Pro  
4  
Planet Audio  
4  
Fugoo  
4  
OmniMount  
3  
GoPro  
3  
Atrend  
3  
G-TECH  
3  
ECO STYLE  
3  
niles  
3  
PreSonus  
3  
Crosley Radio  
3

Peerless Industries  
3  
HP Inc  
3  
x6d inc  
3  
Sabrent  
3  
ZTE  
3  
Beats  
3  
Slingbox  
3  
Case Logic  
2  
Alpine Electronics USA  
2  
Definitive Technology  
2  
Chief Manufacturing  
2  
NVIDIA  
2  
Niles  
2  
Razer USA  
2  
Polk Audio  
2  
1-Year  
2  
Kicker  
2  
Motorola  
2  
Sunpak  
2  
Cooks International LLC (Former Coshell Holdings LLC)  
2  
Energizer  
1  
Diamond Multimedia  
1  
Lenovo  
1  
Lorex  
1  
AUDIOQUEST  
1  
Sonax  
1  
Name: manufacturer, dtype: int64

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

```
In [60]: manufacturer_map=dt_electronics['manufacturer'].value_counts().to_dict()
manufacturer_map
```

```
Out[60]: {'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,
          'Garmin': 24,
          'Parts': 24,
          'Western Digital': 24,
          'YAMAHA': 23,
          'Canon': 23,
          'Logitech': 23,
          'Samsung Electronics': 22,
          'Dell': 21,
          'NETGEAR': 21,
          'Zoom': 21,
          'Panasonic': 20,
          'Denaq': 20,
          'ONKYO': 20,
          'Acer': 19,
          'House of Marley': 18,
          'Power Acoustik': 18,
          'Manufacturers Limited Warranty:2 Year': 18,
          'SiriusXM': 18,
          '120': 18,
          'Yamaha Electronics': 17,
          'Joby': 16,
          'Kanto': 15,
          'Manufacturers Limited Warranty:3 Year': 15,
          'iSimple': 15,
          'Samsung IT': 15,
          'Grace Digital': 15,
          'Asus': 14,
          'Sylvania': 14,
          'Sanus': 13,
          'StarTech.com': 13,
          'Manufacturers Limited Warranty:Lifetime': 13,
          'SiliconDust USA, Inc': 13,
          'Bose': 13,
          'SanDisk': 13,
          'Razer': 13,
          'Papago': 13,
          'MTX': 13,
          'Manfrotto': 13,
          'Grace Digital Audio': 13,
```

'Logitech, Inc': 12,  
'Lenovo Corporate': 12,  
'Belkin': 12,  
'Vizio': 12,  
'JVC': 12,  
'Kenwood': 12,  
'Audioquest': 11,  
'TP-Link': 11,  
'Manufacturers Limited Warranty:1 Year': 11,  
'Boytone': 11,  
'OtterBox': 11,  
'Ecoxgear': 11,  
'BIC America': 11,  
'iHome': 11,  
'RAZER': 11,  
'Jensen': 11,  
'Westone Laboratories, Inc.': 11,  
'NZXT': 11,  
'Sony Mobile Communications, (USA) Inc': 10,  
'RCA': 10,  
'House Of Marley': 10,  
'SAMSUNG': 10,  
'LG ODD': 10,  
'Sima': 10,  
'ISIMPLE': 10,  
'KICKER': 10,  
'Audiovox': 9,  
'Canon Cameras US': 9,  
'Netgear Inc': 9,  
'RODE': 9,  
'Sherwood': 9,  
'JAYBIRD': 8,  
'Dual': 8,  
'Zagg': 8,  
'1 year warranty': 8,  
'Jbl Everest 110 In-Ear Wireless Bluetooth Headphones (silver)': 8,  
'Tiffen': 8,  
'AudioQuest': 8,  
'Midland': 8,  
'Sharp': 8,  
'Sigma': 7,  
'Cybertronpc': 7,  
'JBL': 7,  
'Pyle': 7,  
'Spacepole Ince': 7,  
'MEE audio': 7,  
'Peerless': 7,  
'PYLE': 7,  
'Apple Computer': 7,  
'Visiontek': 7,  
'Viper by GLD Products': 7,  
'Microsoft': 7,  
'Peak Design': 7,  
'Lenovo Idea': 7,  
'5 Years': 6,  
'ELITE SCREENS DIRECTSHIP': 6,  
'LowePro': 6,

'Speck': 6,  
'Elite Screens': 6,  
'ARRIS Solutions, Inc.': 6,  
'SKB Cases': 6,  
'1 Year': 6,  
'Allround Software': 6,  
'Apple Computer (Direct)': 6,  
'Klipsch': 6,  
'Coolermaster': 6,  
'Zmodo': 6,  
'MTRLC LLC': 6,  
'BOSE': 6,  
'Pelican': 6,  
'Aluratek': 6,  
'Spartan': 6,  
'Limited Lifetime Warranty': 6,  
'Polaroid': 6,  
'KLIPSCH': 6,  
'Bower Camera': 5,  
'eTape16': 5,  
'Marantz': 5,  
'Buffalo Technology': 5,  
'Incipio': 5,  
'Sonos': 5,  
'Novatel': 5,  
'Ilive Blue': 5,  
'Boyo': 5,  
'TiVo': 5,  
'Urban Armor Gear': 5,  
'Sound Around': 5,  
'Cerwin Vega': 4,  
'Jvc Real Sound System Z Series Ha-Sz2000 Stream Woofer Db Headphones (black, Japan Import) V-Moda': 4,  
'Pac': 4,  
'Klipsch X12 Bluetooth Neckband Headphones (black Leather)': 4,  
'Fusion': 4,  
'Siriusxm': 4,  
'StarTech': 4,  
'Bower': 4,  
'Emerge': 4,  
'2 Years': 4,  
'EPSON AMERICA, INC.': 4,  
'Sonic Alert': 4,  
'Thule': 4,  
'Dantona': 4,  
'TRANSCEND Storage': 4,  
'Mophie': 4,  
'Go-Pro': 4,  
'Planet Audio': 4,  
'Fugoo': 4,  
'OmniMount': 3,  
'GoPro': 3,  
'Atrend': 3,  
'G-TECH': 3,  
'ECO STYLE': 3,  
'niles': 3,  
'PreSonus': 3,

```
'Crosley Radio': 3,  
'Peerless Industries': 3,  
'HP Inc': 3,  
'x6d inc': 3,  
'Sabrent': 3,  
'ZTE': 3,  
'Beats': 3,  
'Slingbox': 3,  
'Case Logic': 2,  
'Alpine Electronics USA': 2,  
'Definitive Technology': 2,  
'Chief Manufacturing': 2,  
'NVIDIA': 2,  
'Niles': 2,  
'Razer USA': 2,  
'Polk Audio': 2,  
'1-Year': 2,  
'Kicker': 2,  
'Motorola': 2,  
'Sunpak': 2,  
'Cooks International LLC (Former Coshell Holdings LLC)': 2,  
'Energizer': 1,  
'Diamond Multimedia': 1,  
'Lenovo': 1,  
'Lorex': 1,  
'AUDIOQUEST': 1,  
'Sonax': 1}
```

here I will replace 'Apple' with 101

```
In [61]: dt_electronics['manufacturer_count']=dt_electronics['manufacturer'].map(lambda x: len(x))
dt_electronics.head()
```

Out[61]:

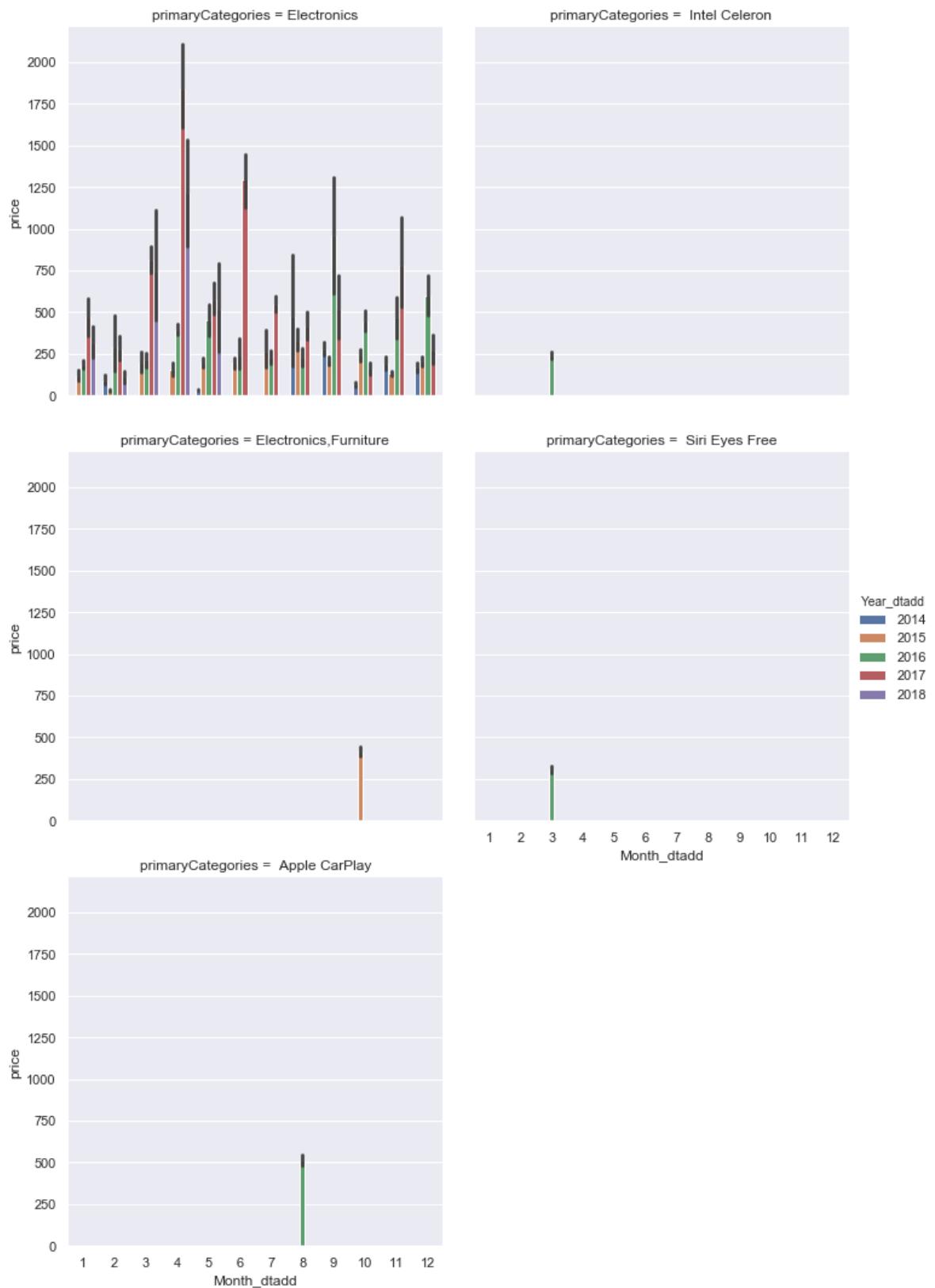
	pricesavailability	pricescondition	pricesdateSeen	pricesmerchant	pricesshipping
0	1590	4645	2017-05-05T20:00:00Z,2017-05-09T15:00:00Z	2109	
1	1590	4645	2017-10-08T02:00:00Z,2017-12T03:00:00Z,2017...	2109	
2	1590	4645	2017-10-09T19:00:00Z,2017-12T14:00:00Z,2017...	2109	
3	1590	4645	2017-09-09T05:00:00Z,2017-18T13:00:00Z,2017...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

```
In [62]: dt_electronics.columns
```

```
Out[62]: 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_update
d',
       'Year_updated', 'manufacturer_count'],
      dtype='object')
```

I will plot primaryCategories with dateAdded

```
In [63]: 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')
```



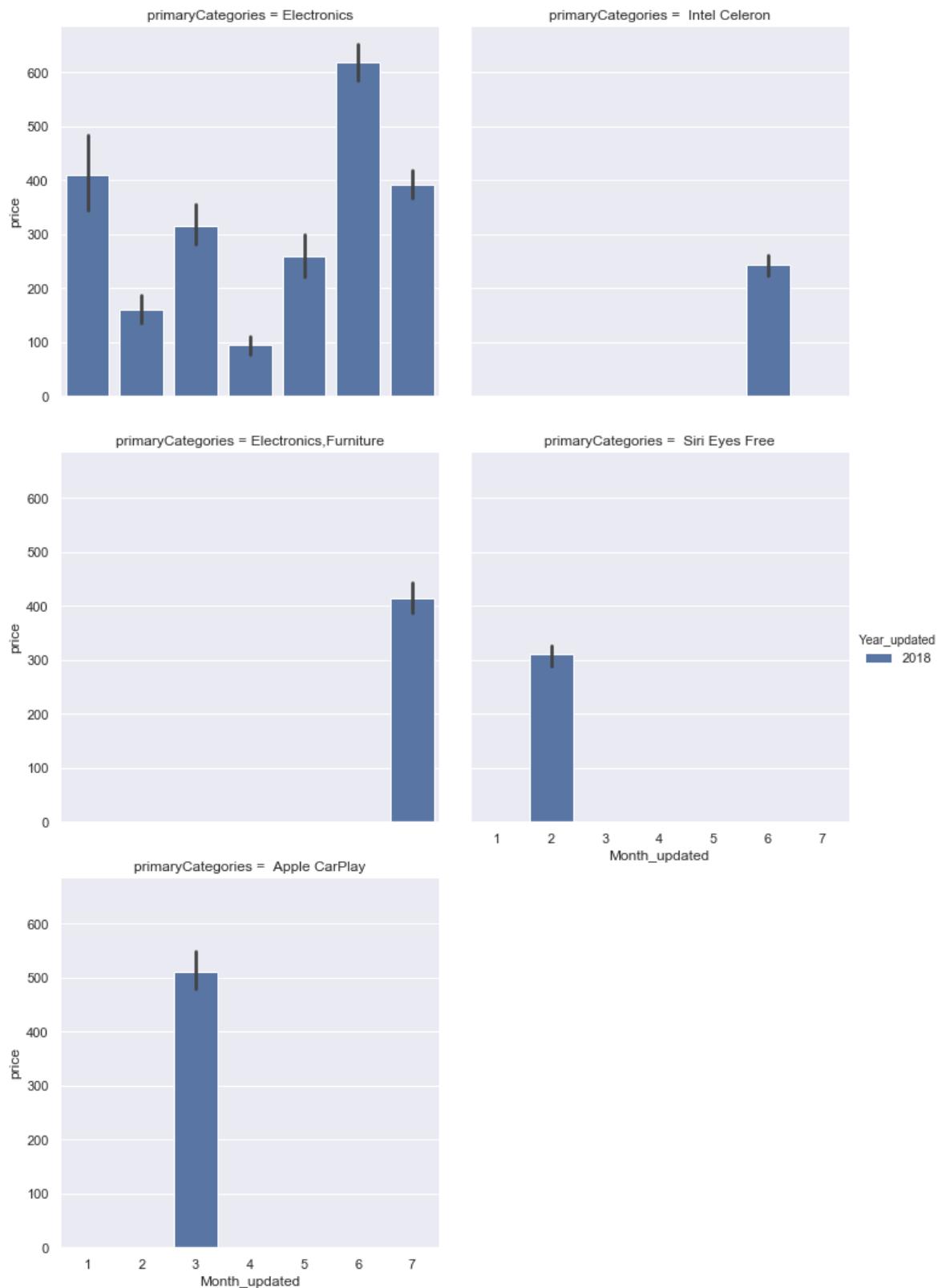
## **Observation**

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 [64]: 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')
```



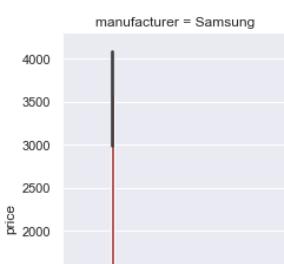
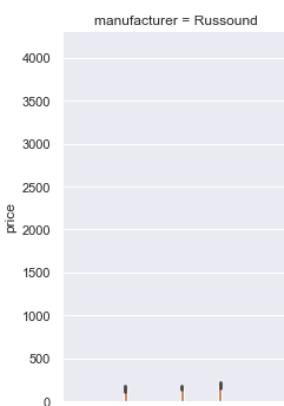
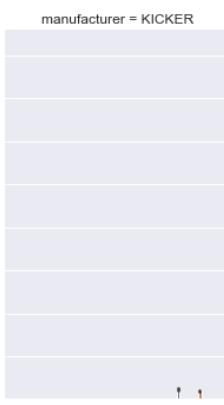
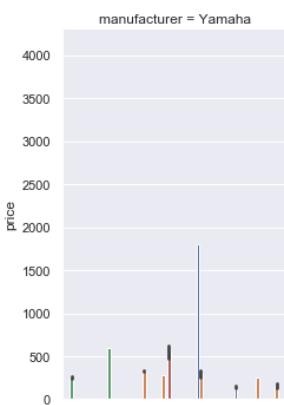
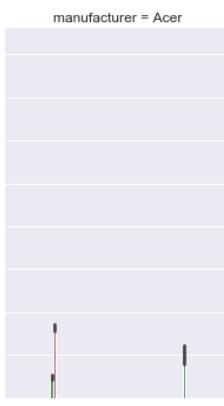
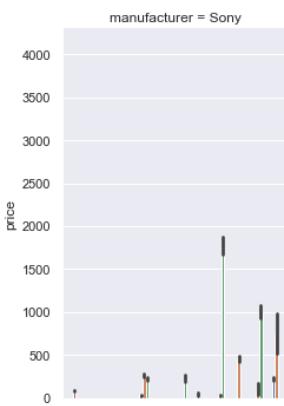
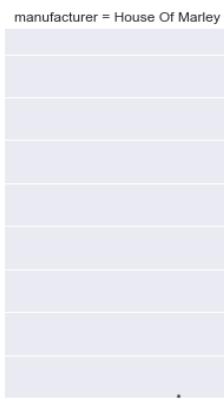
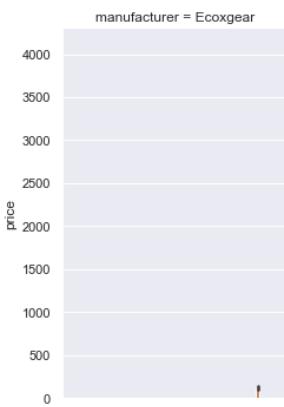
## **Observation**

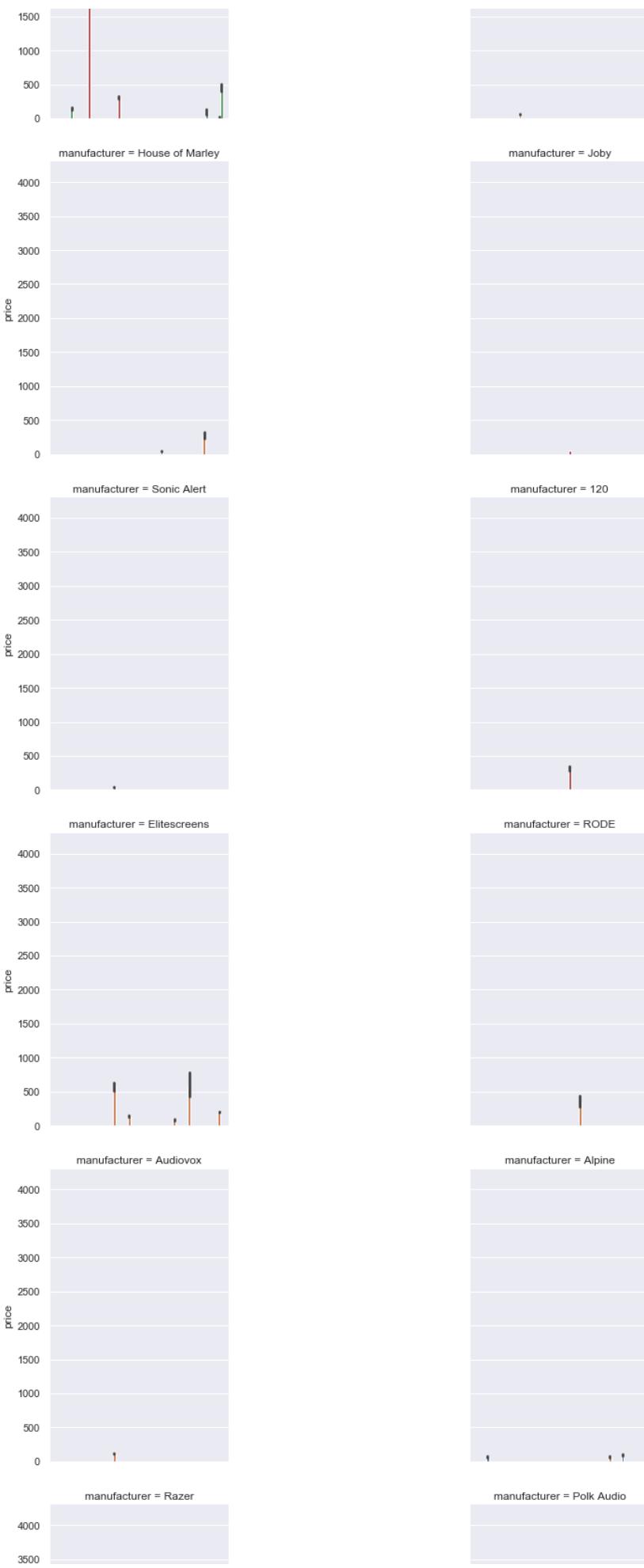
From this graph we can conclude Electronics category has more price than other category

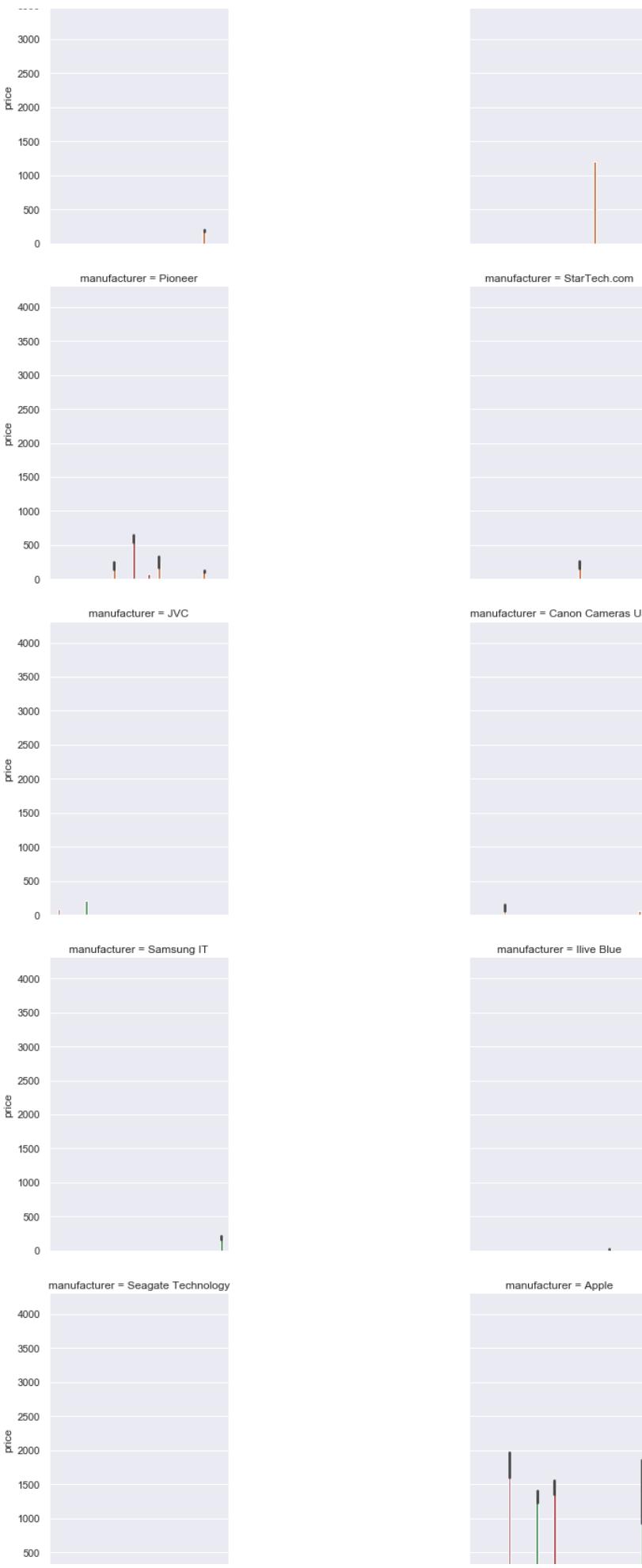
I can also conclude that in 6th month of 2018 price is highest

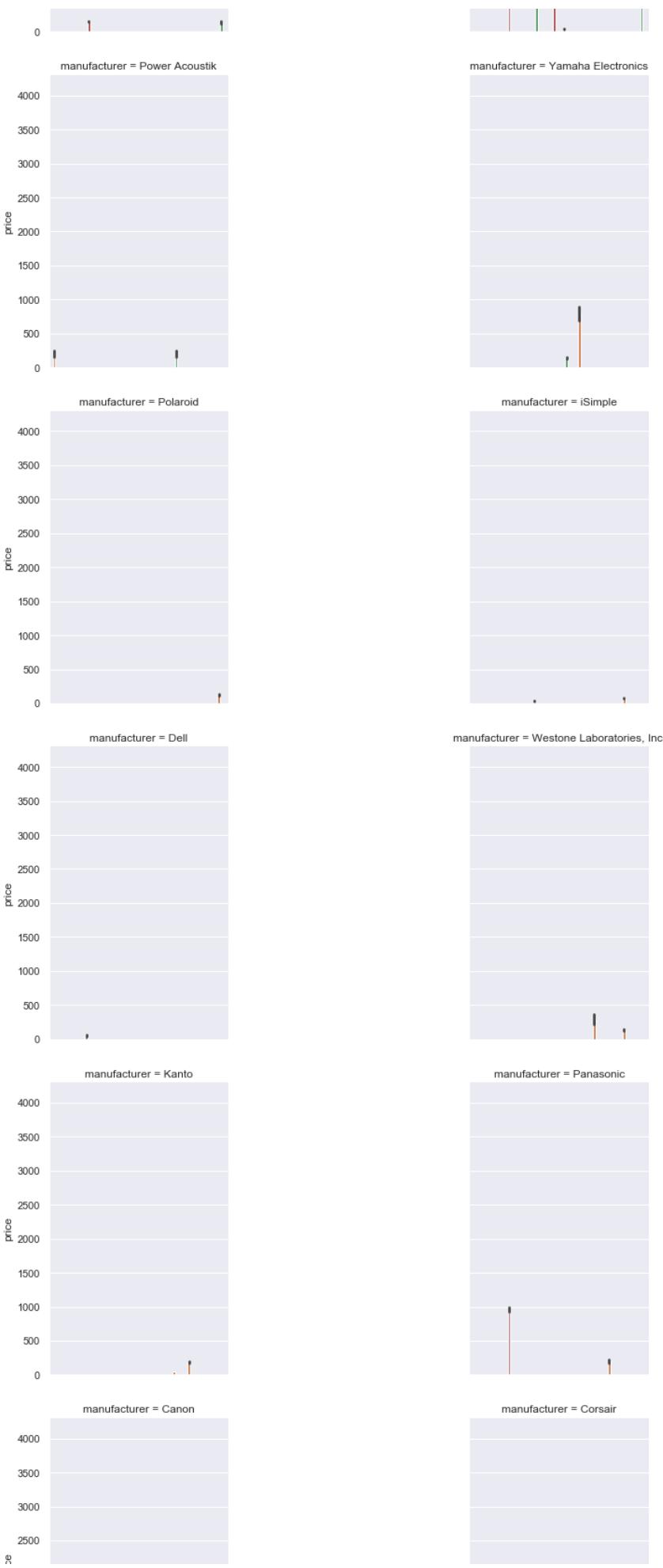
I will plot manufacturer with dateAdded

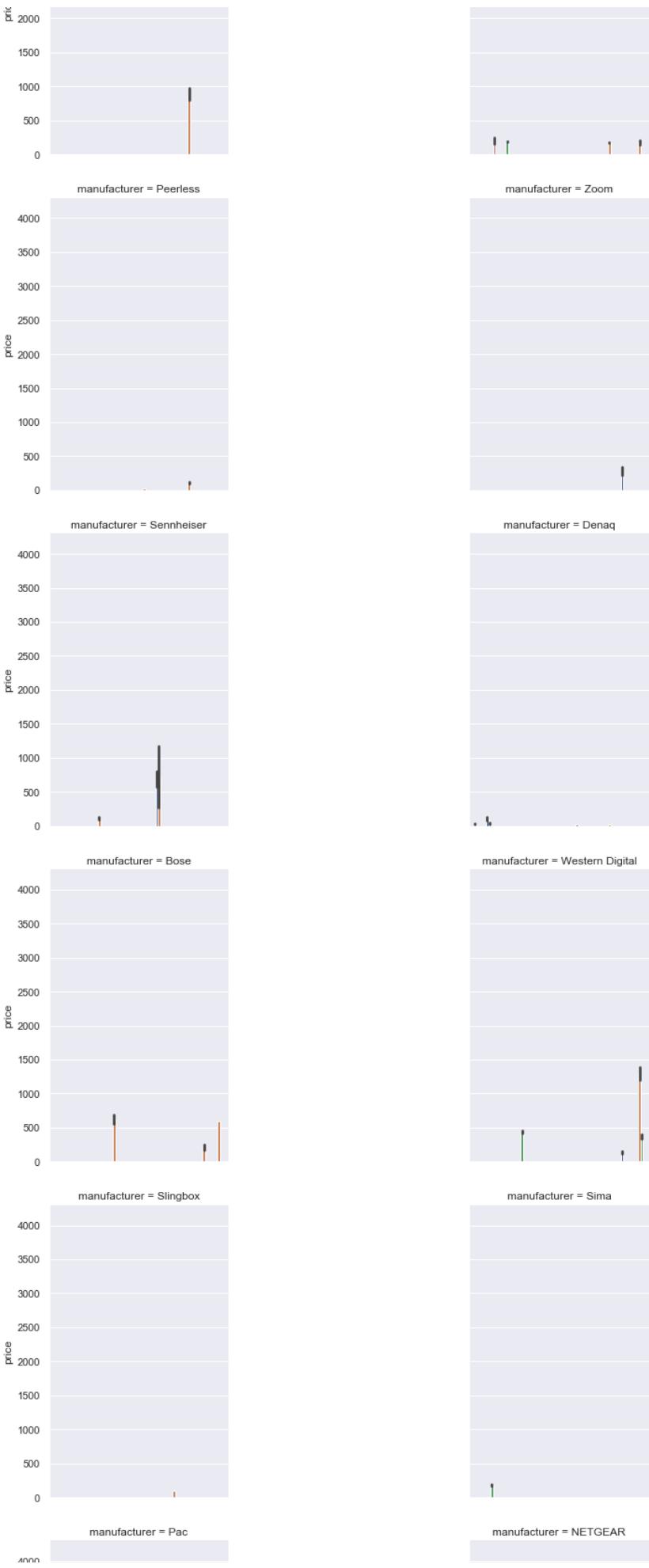
```
In [65]: sns.set(style='darkgrid', font_scale=1.0)
sns.catplot(x='Month_dtadd',
            y='price',hue='Year_dtadd',
            col='manufacturer',
            col_wrap=2,
            data=dt_electronics,
            kind='bar')
plt.savefig('Dateadded_manufacture.png')
```

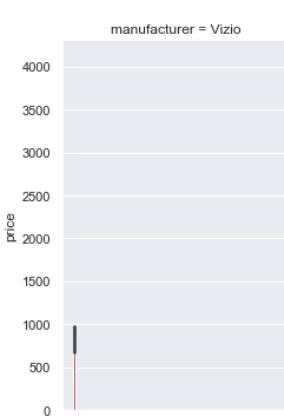
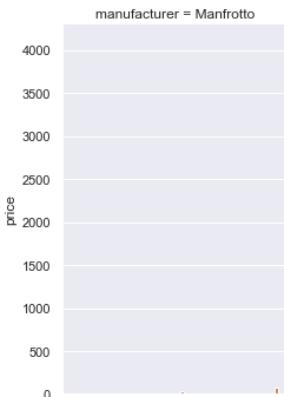
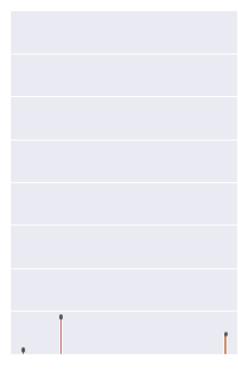


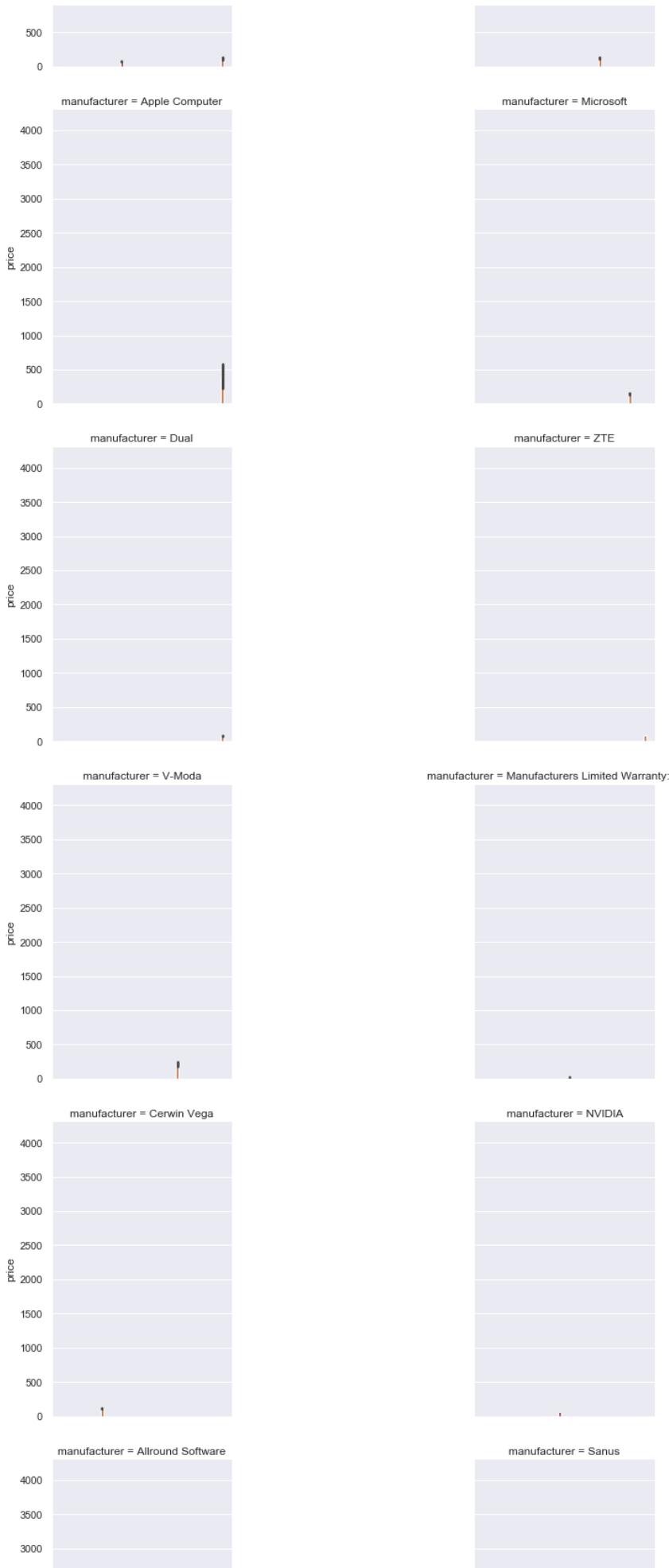






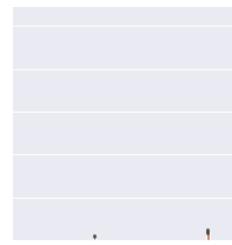








manufacturer = Samsung Electronics

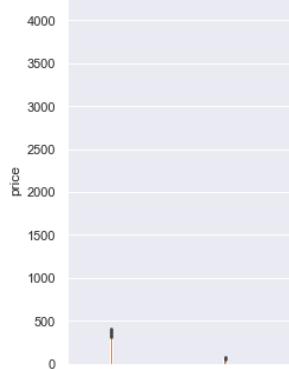


manufacturer = Definitive Technology



manufacturer = Parts

manufacturer = Logitech



manufacturer = Asus

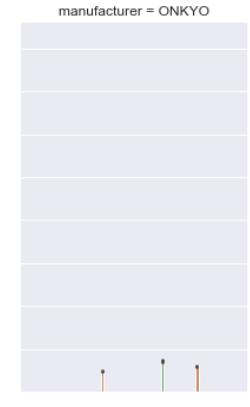
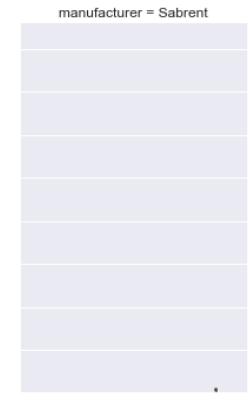
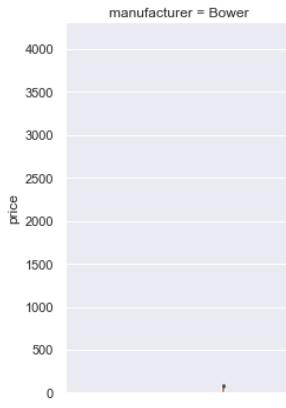
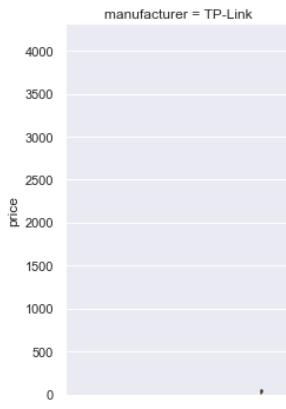
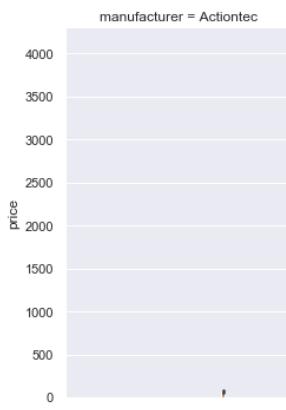
manufacturer = Lowepro

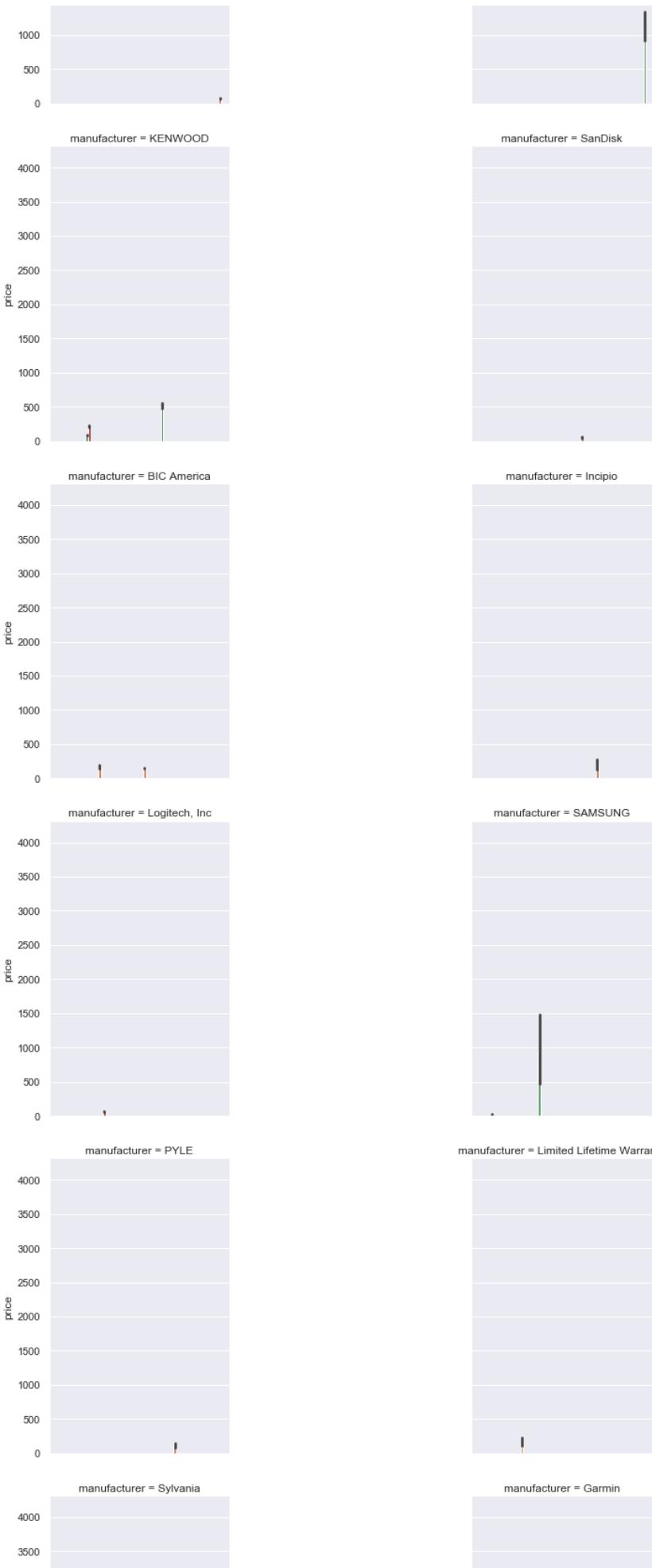


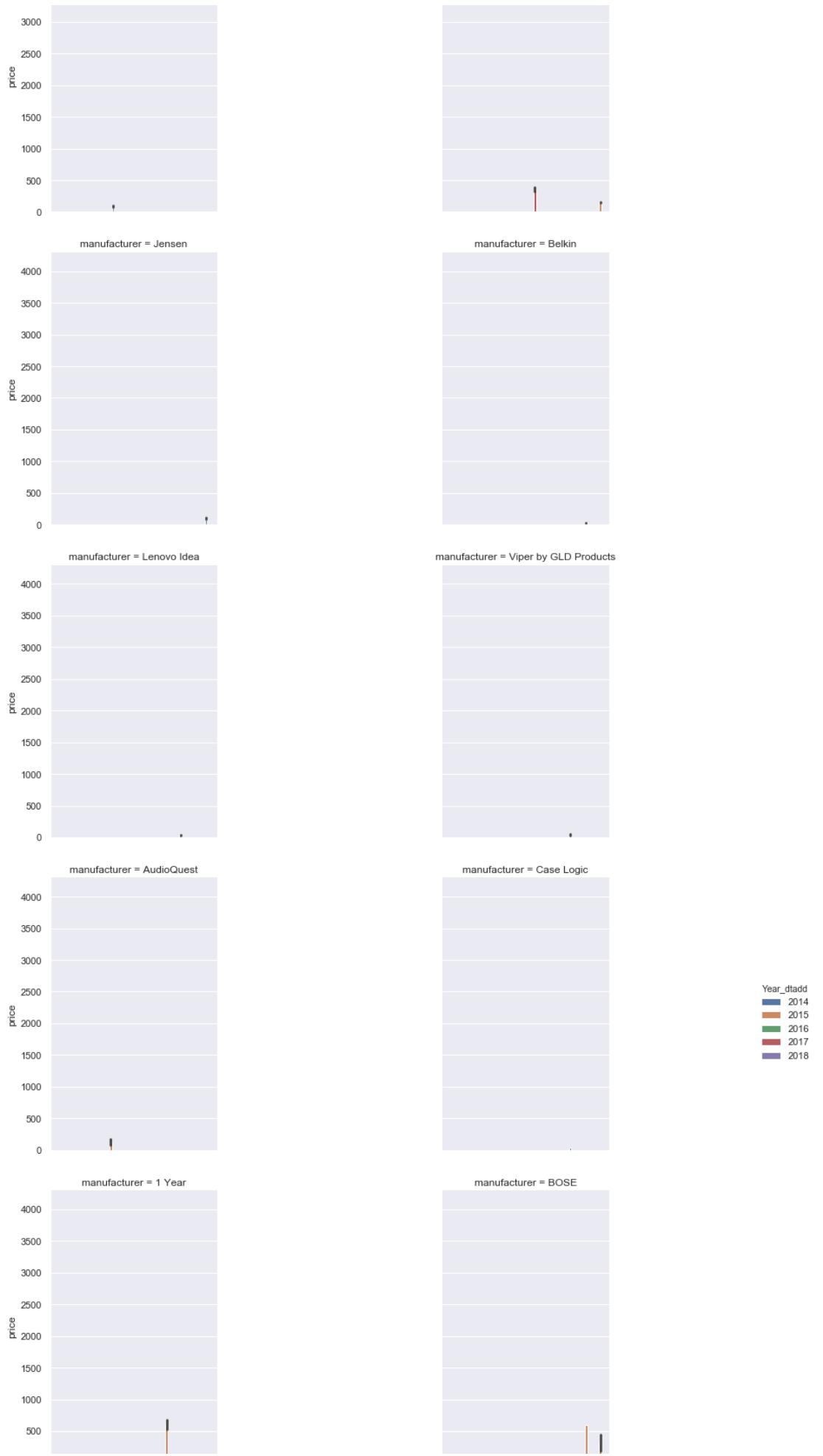
manufacturer = G-TECH

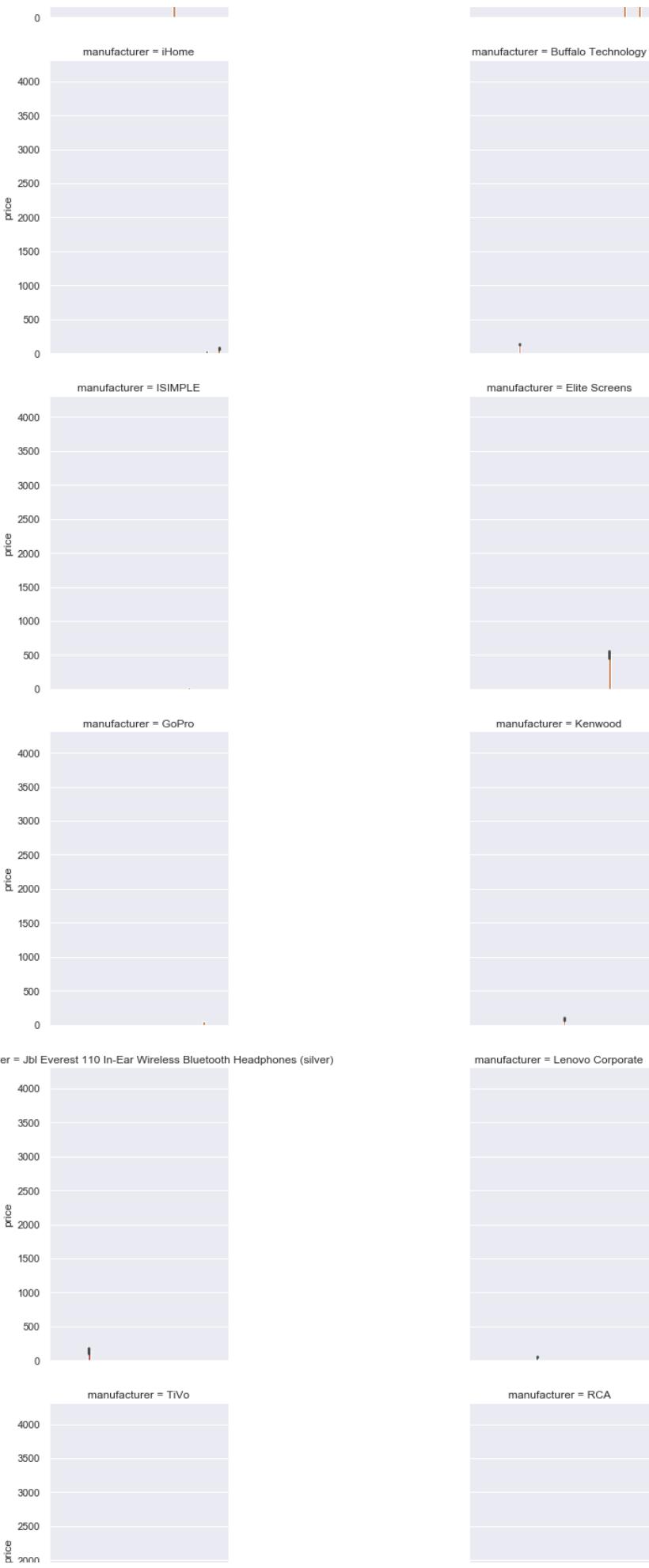
manufacturer = Denon

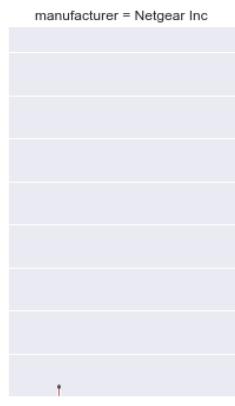
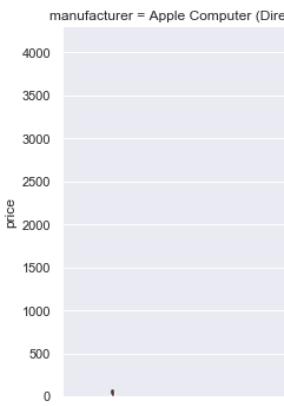
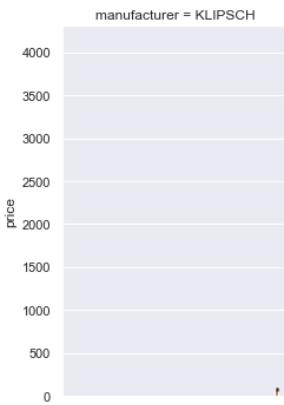
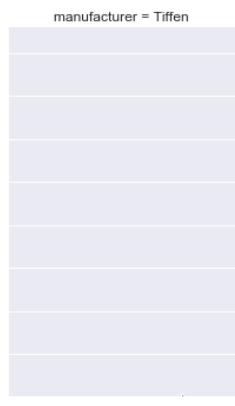
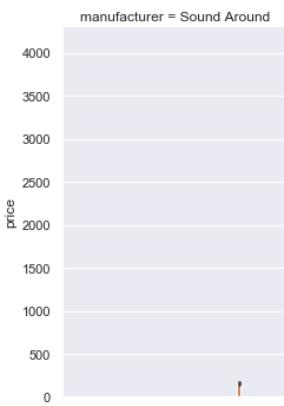
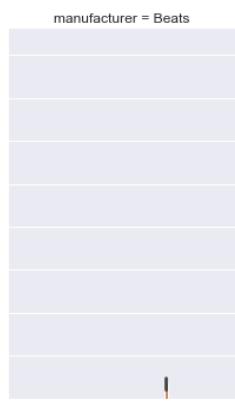
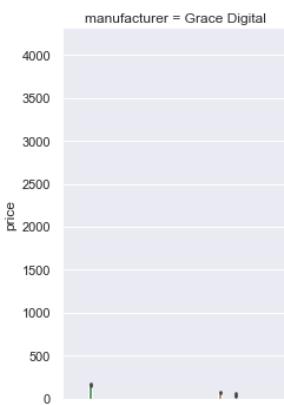


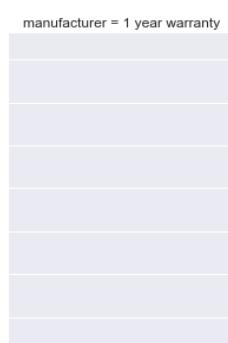
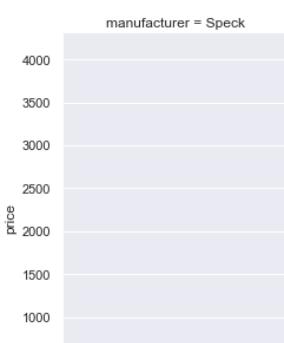
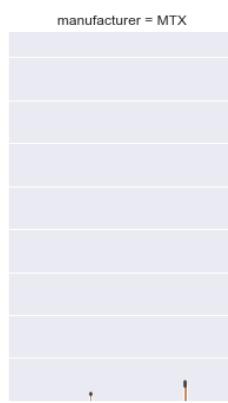
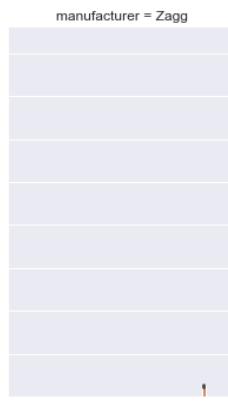
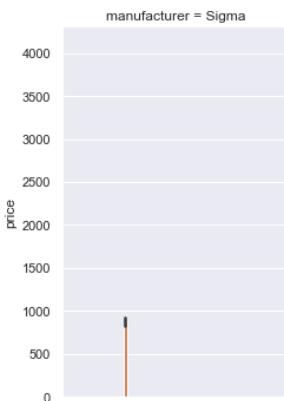
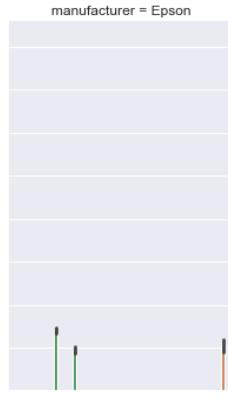
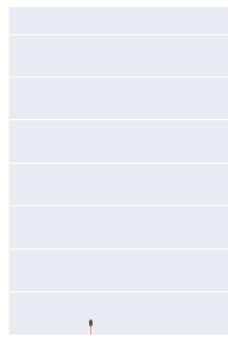
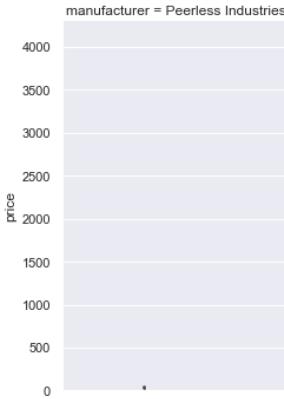
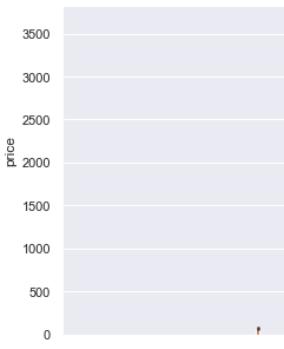


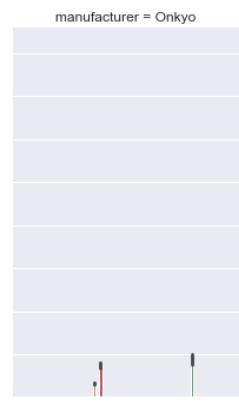
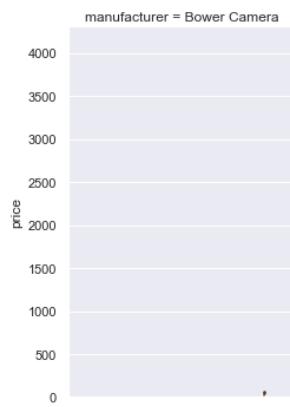
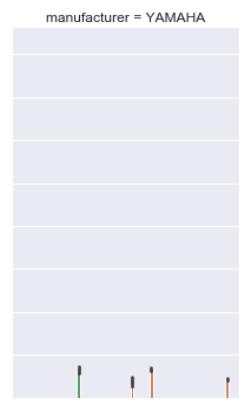
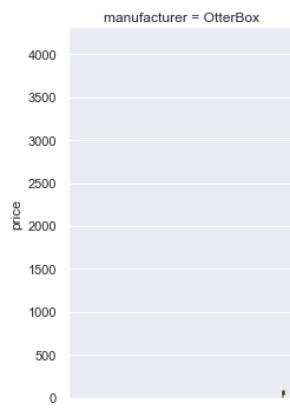
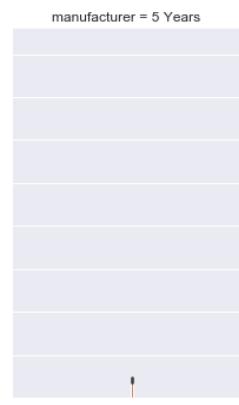
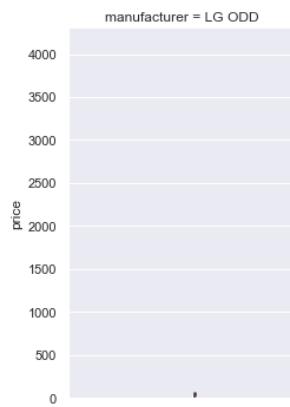
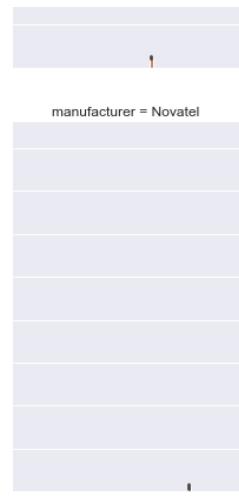
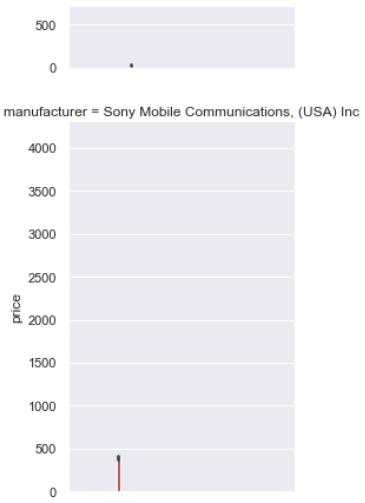


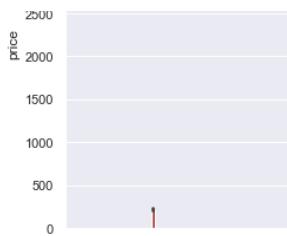




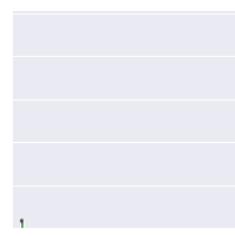




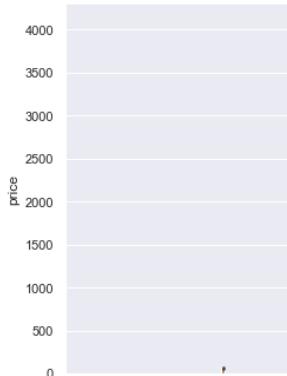




manufacturer = Pelican



manufacturer = EPSON AMERICA, INC.

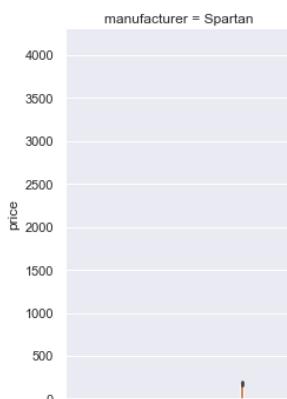


manufacturer = Visiontek

manufacturer = Audioquest



manufacturer = Audioquest



manufacturer = Spartan

manufacturer = RAZER



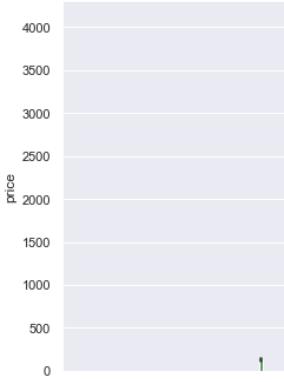
manufacturer = Boyo



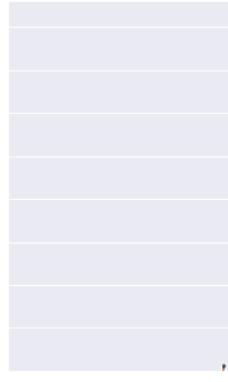
manufacturer = TRANSCEND Storage

manufacturer = JAYBIRD

manufacturer = Urban Armor Gear



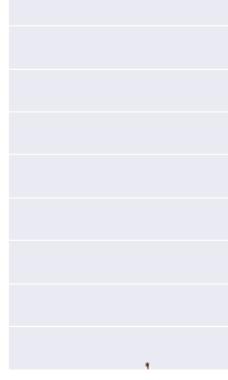
manufacturer = NZXT



manufacturer = SKB Cases



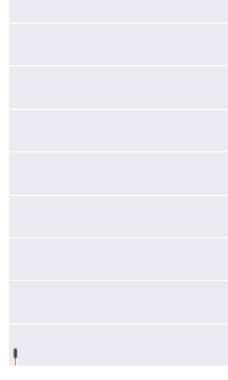
manufacturer = Planet Audio



manufacturer = Fugoo



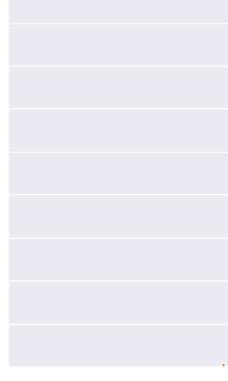
manufacturer = PreSonus



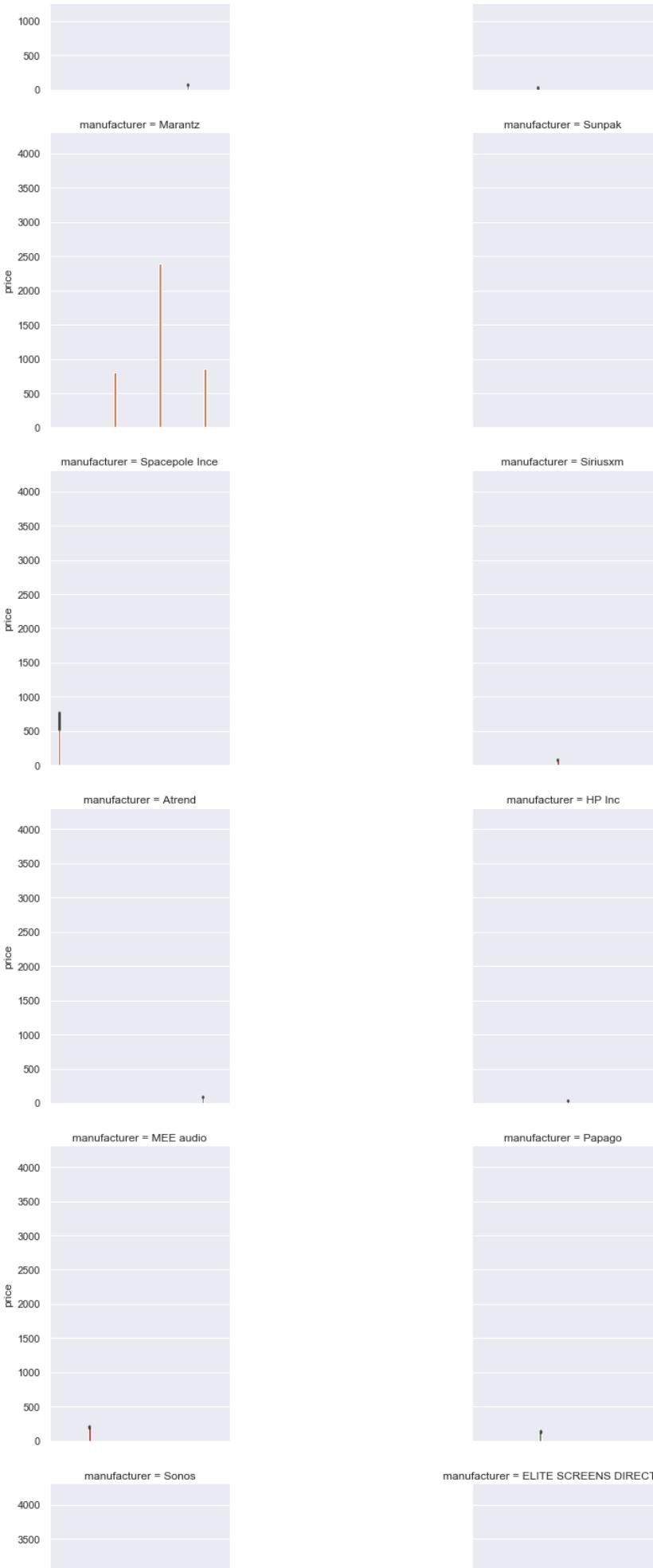
manufacturer = Peak Design

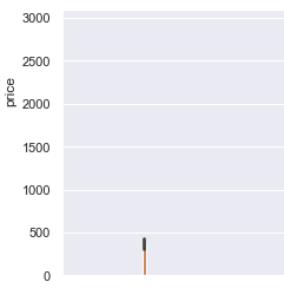


manufacturer = Fusion

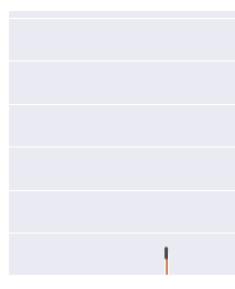


manufacturer = Go-Pro





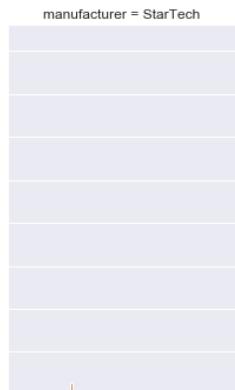
manufacturer = AUDIOQUEST



manufacturer = x6d inc



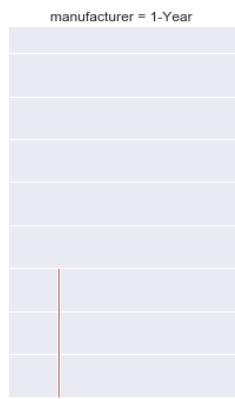
manufacturer = ECO STYLE



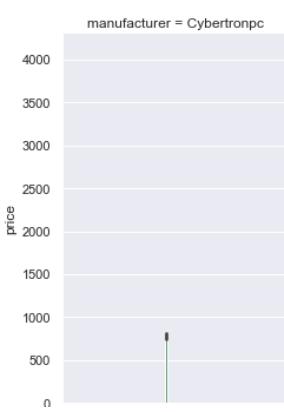
manufacturer = StarTech



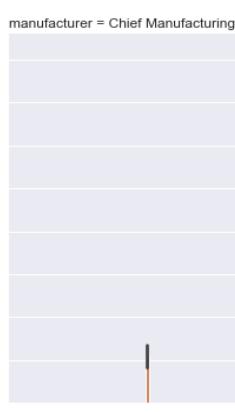
manufacturer = Kicker



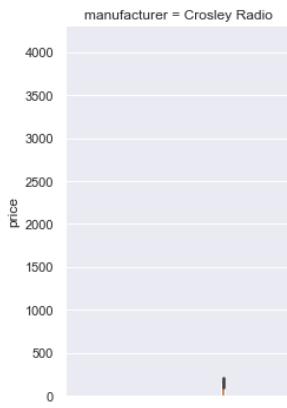
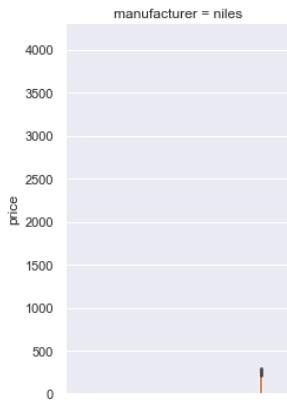
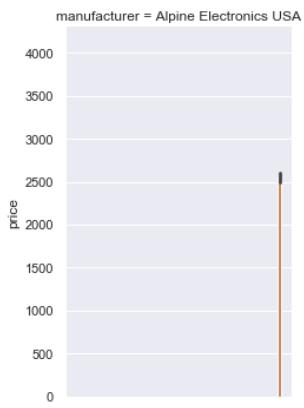
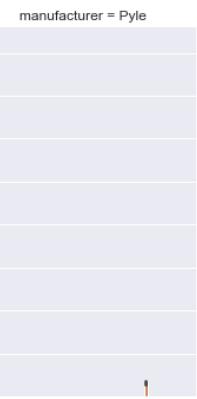
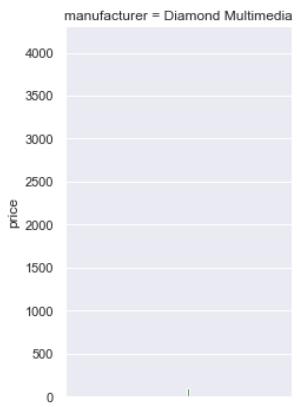
manufacturer = 1-Year

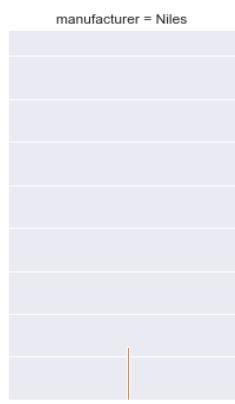
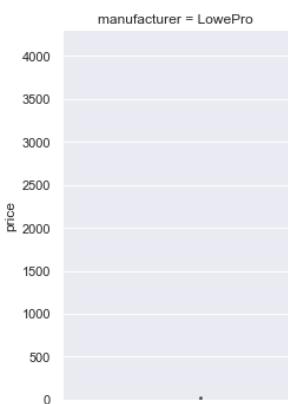
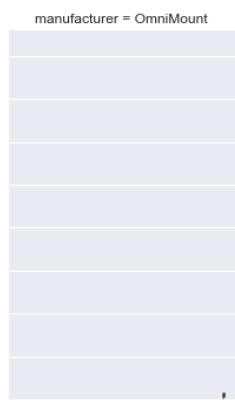
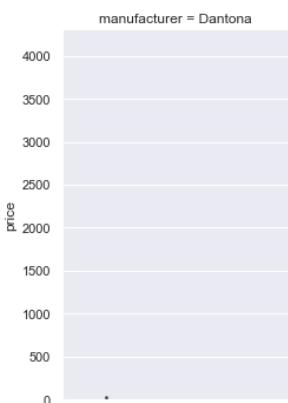
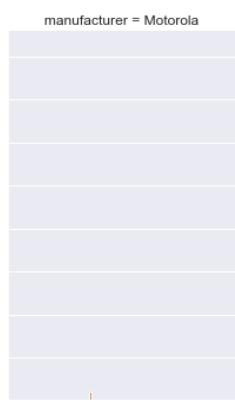
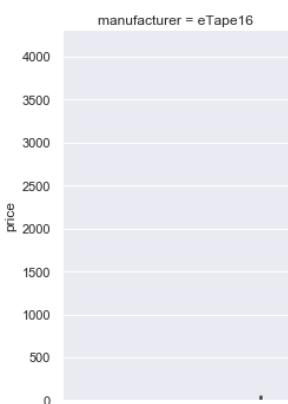
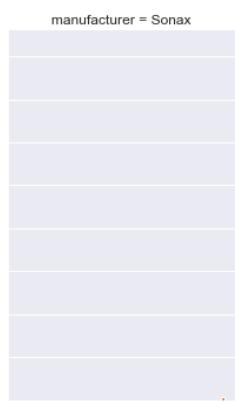
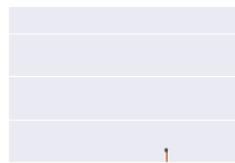
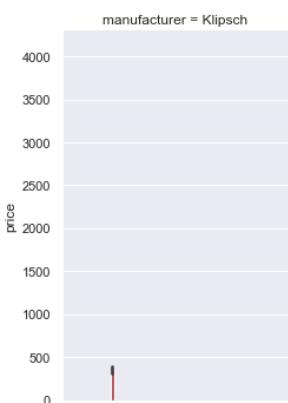
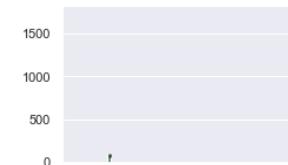


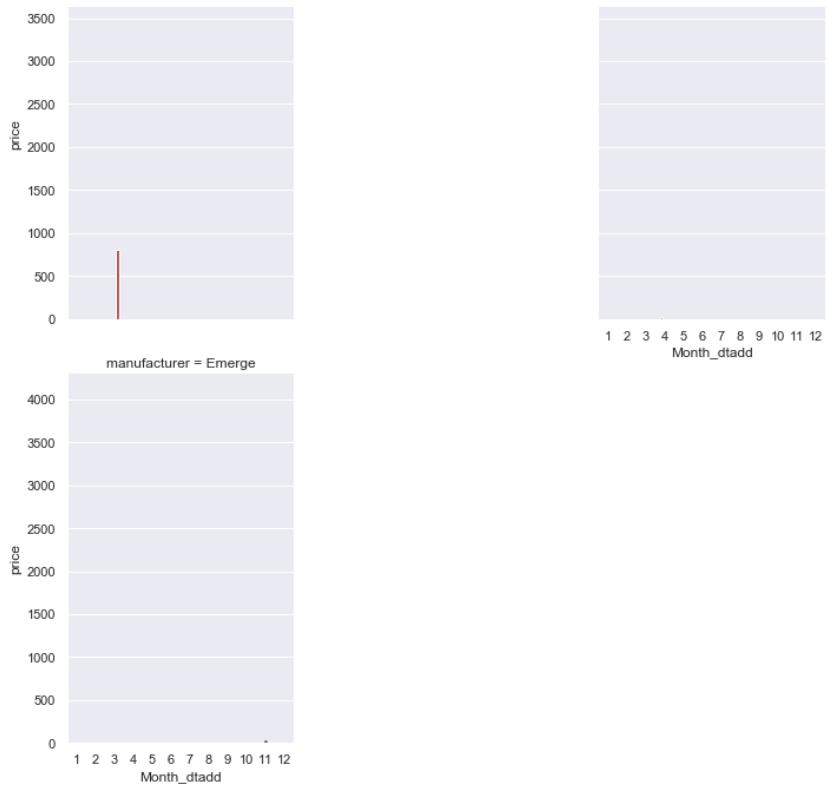
manufacturer = Cybertronpc



manufacturer = Chief Manufacturing





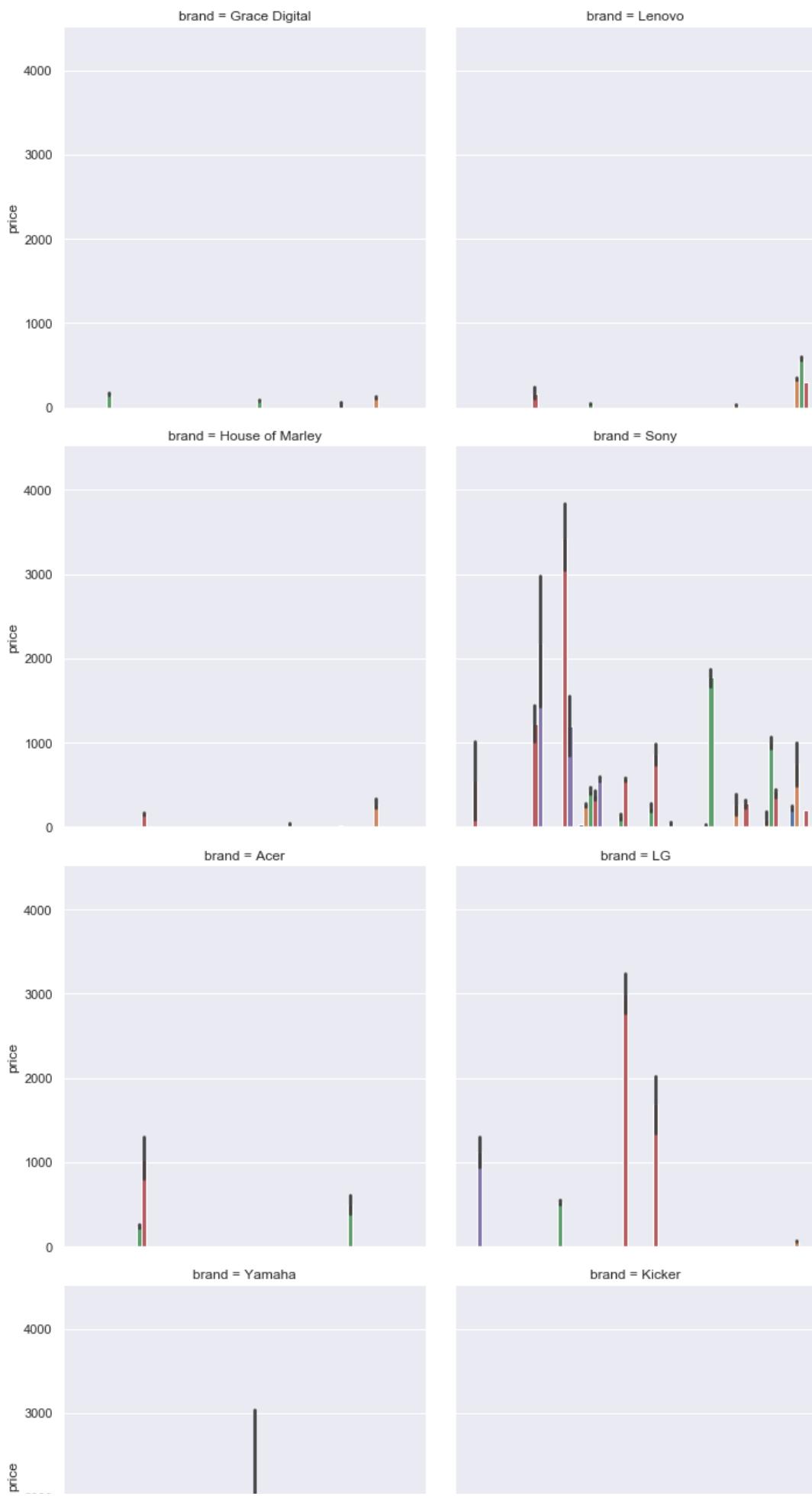


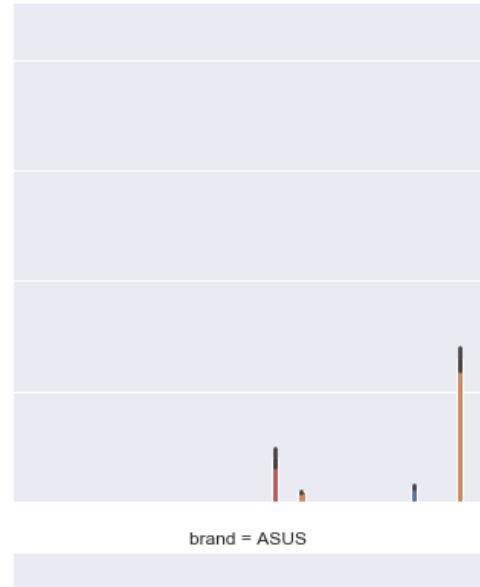
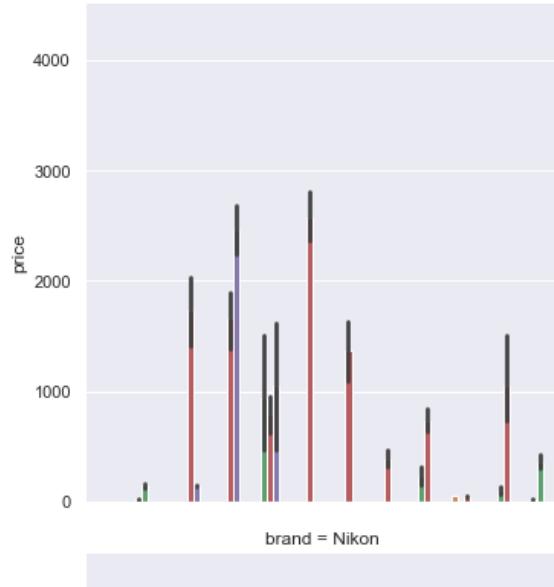
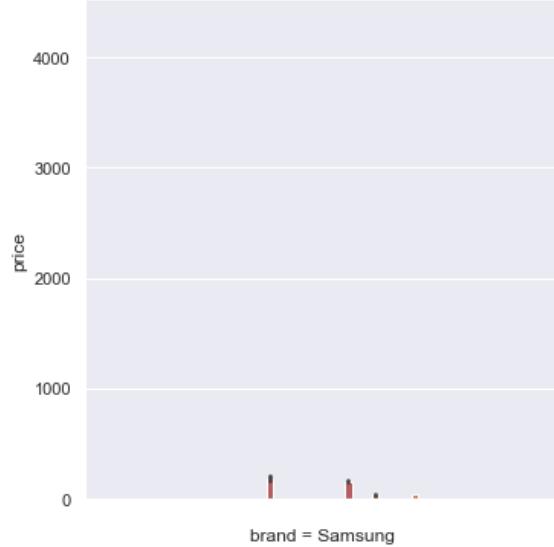
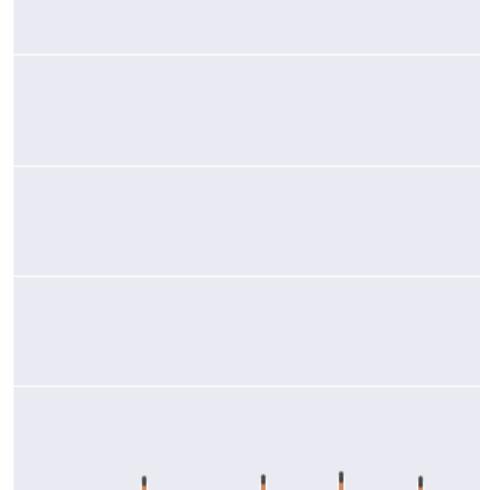
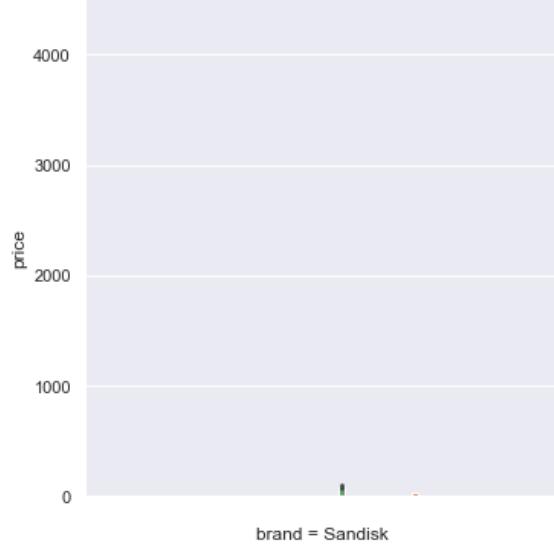
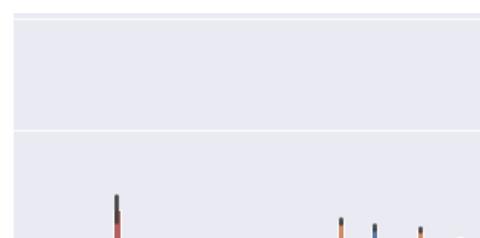
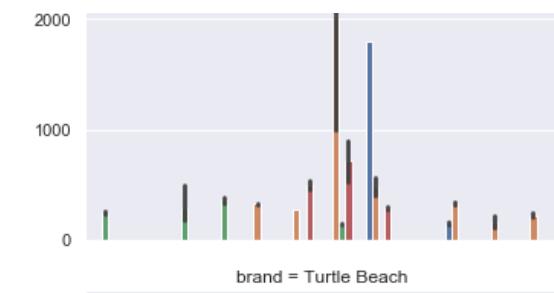
## Observation

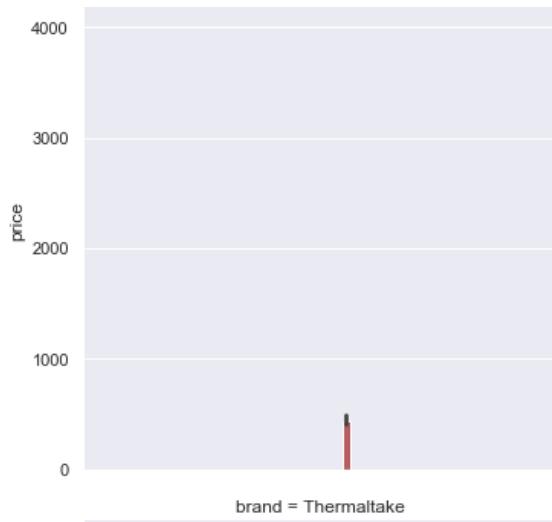
From this graph I can conclude Samsung has more price than other category  
I can also conclude that in 3rd month of 2017 price is highest

I will plot brand with dateAdded

```
In [66]: 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')
```



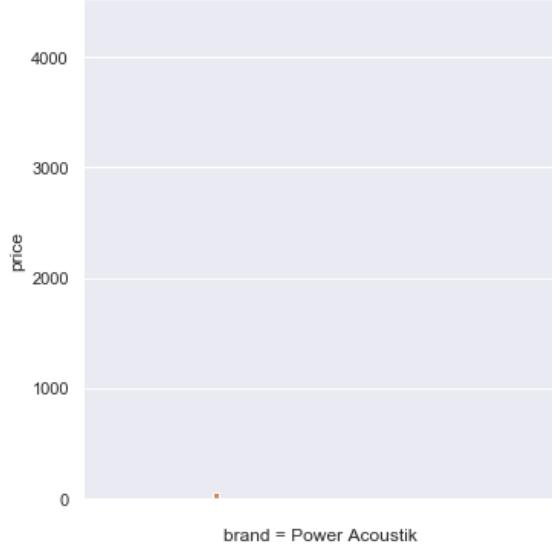




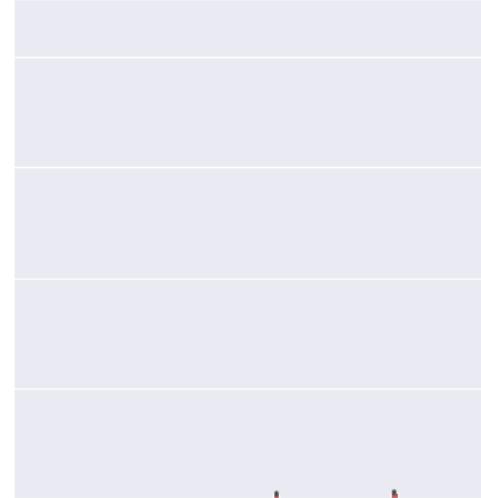
brand = Thermaltake



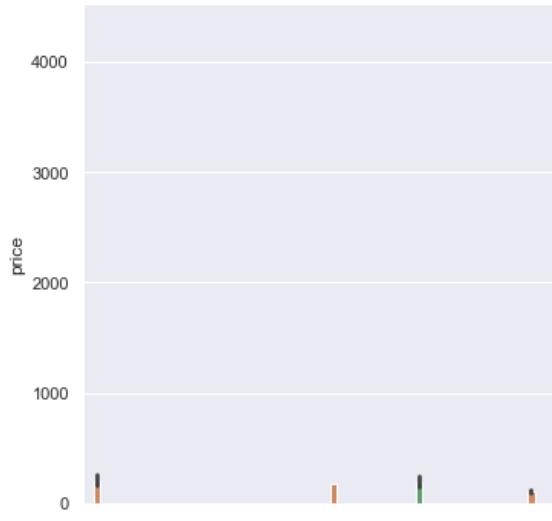
brand = Joby



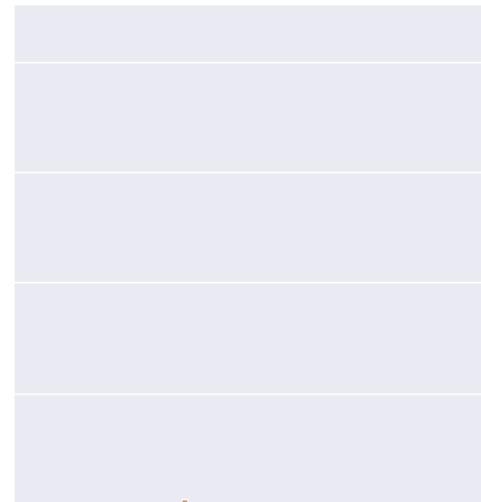
brand = Power Acoustik



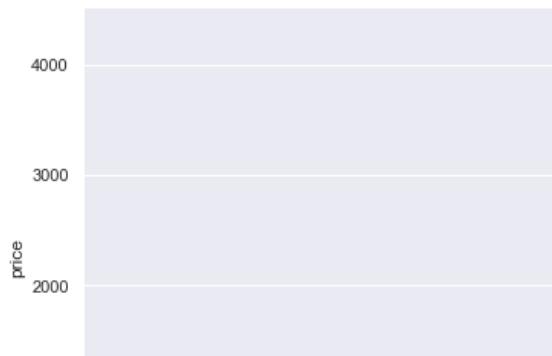
brand = Sonic Alert

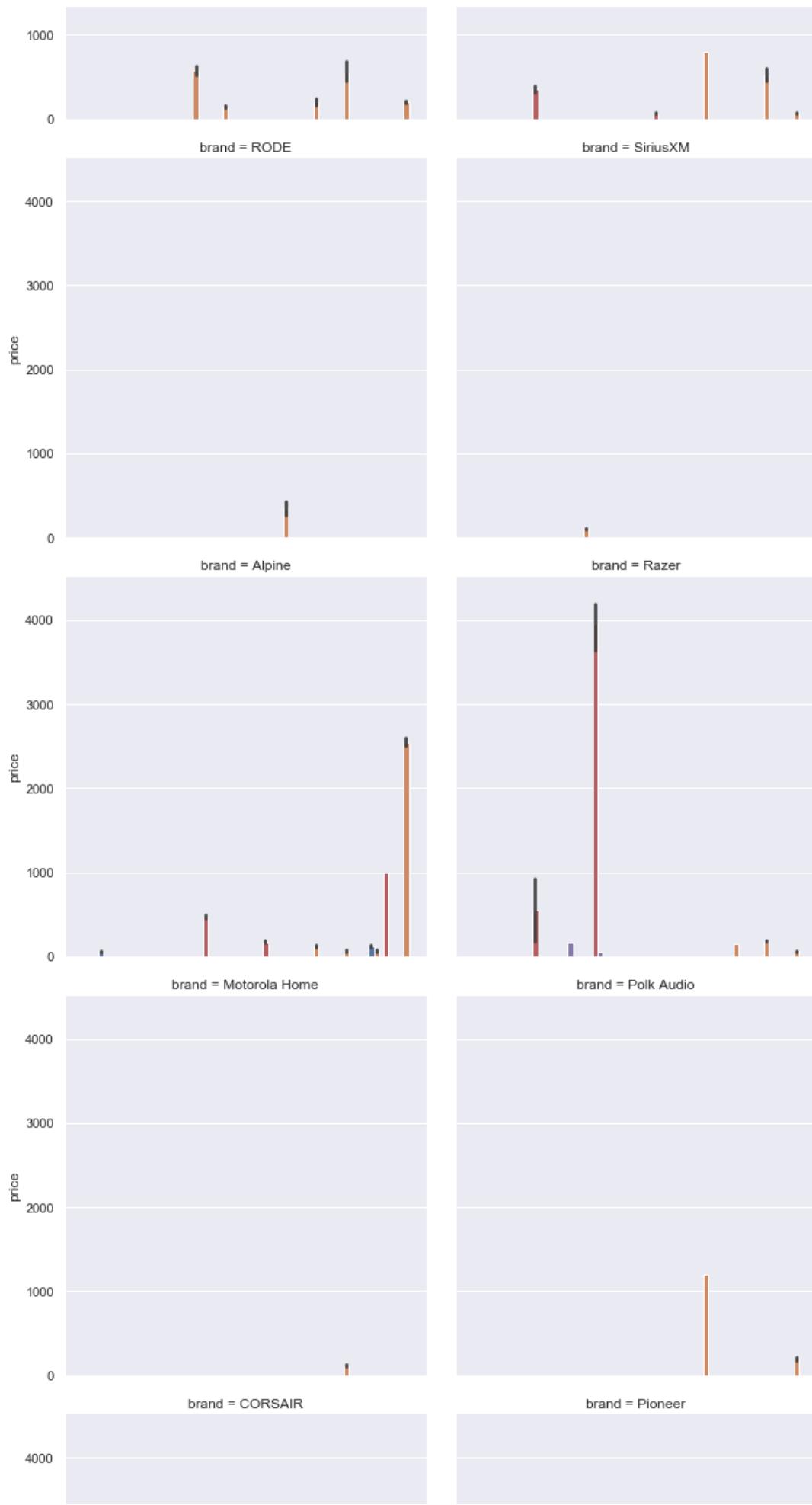


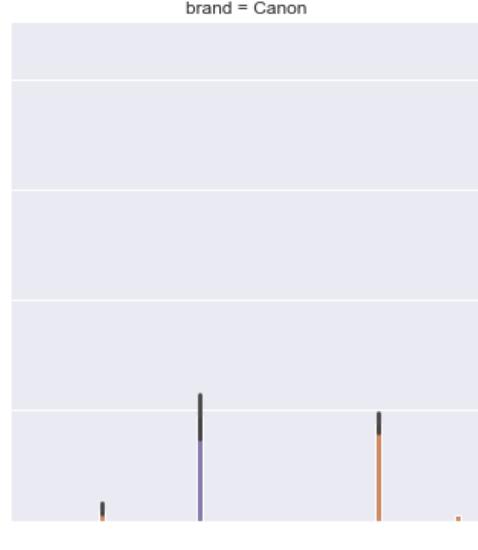
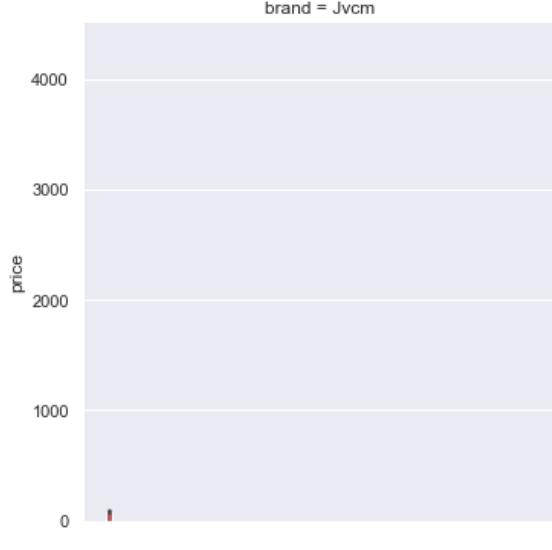
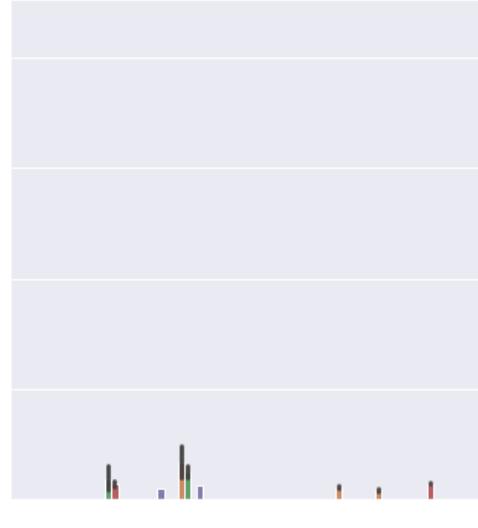
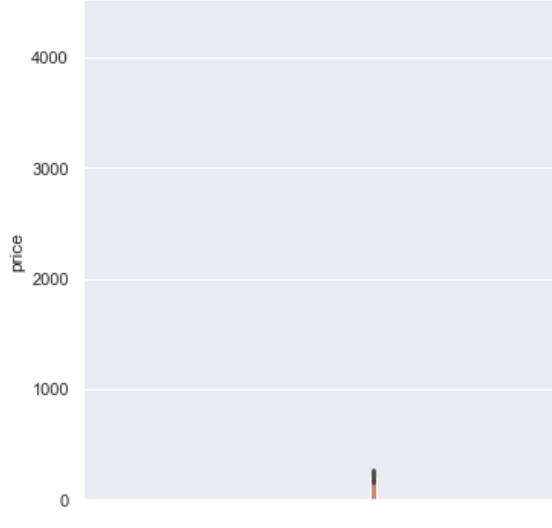
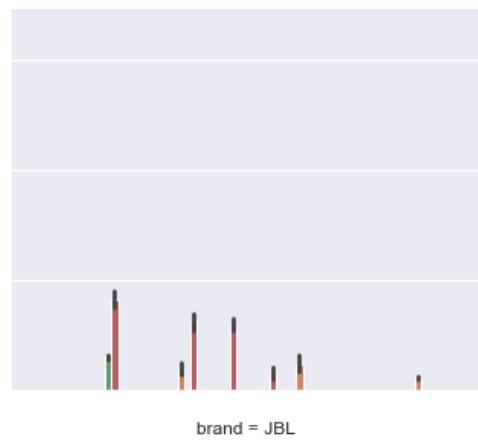
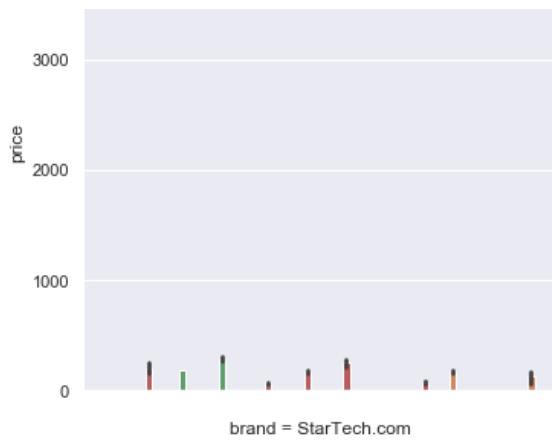
brand = Elite Screens

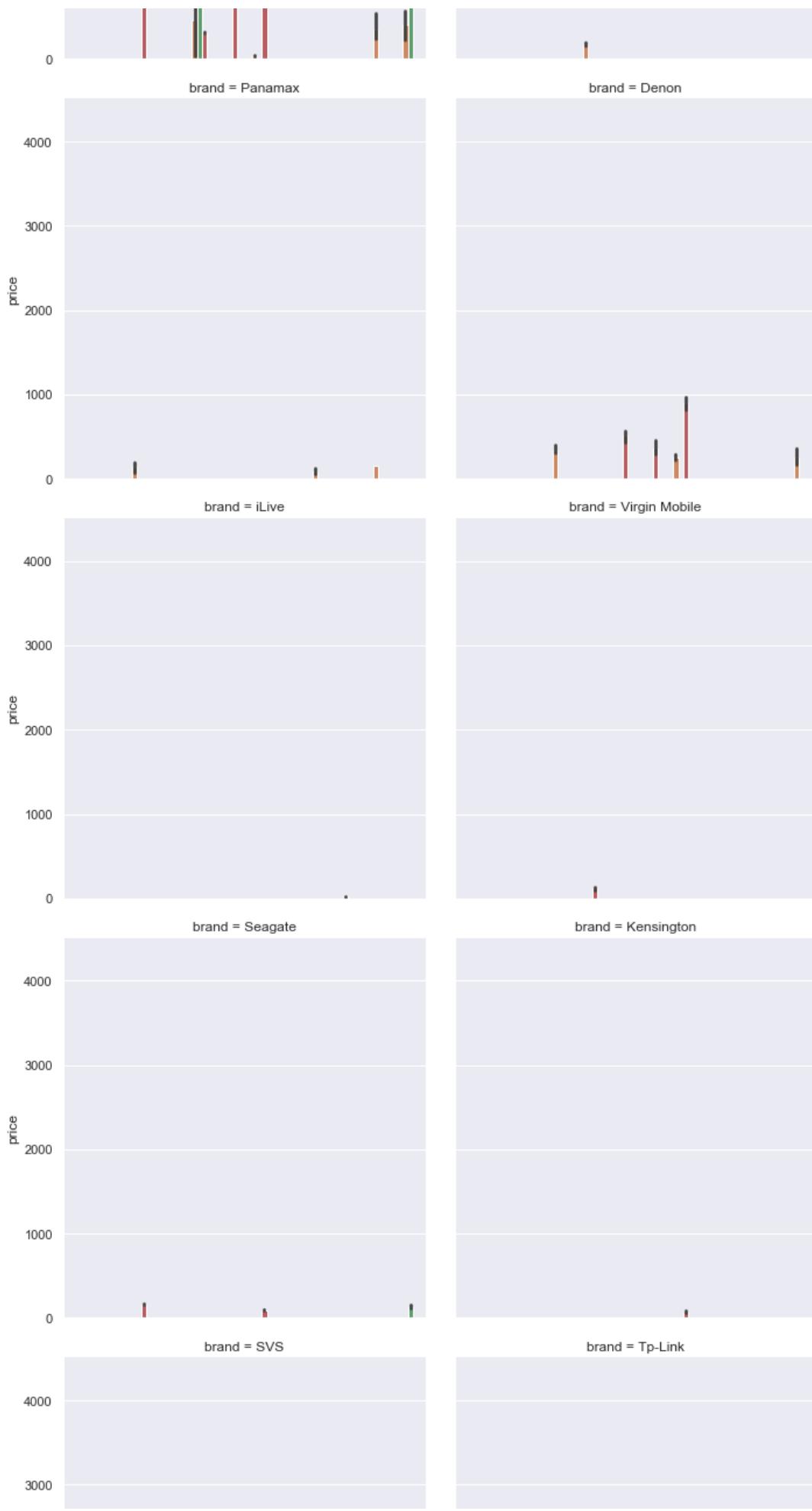


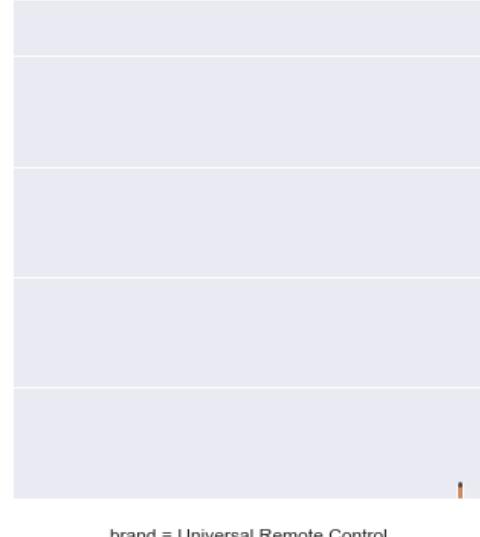
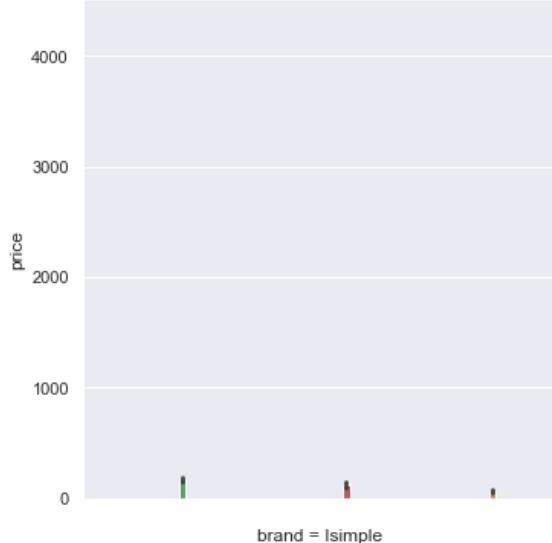
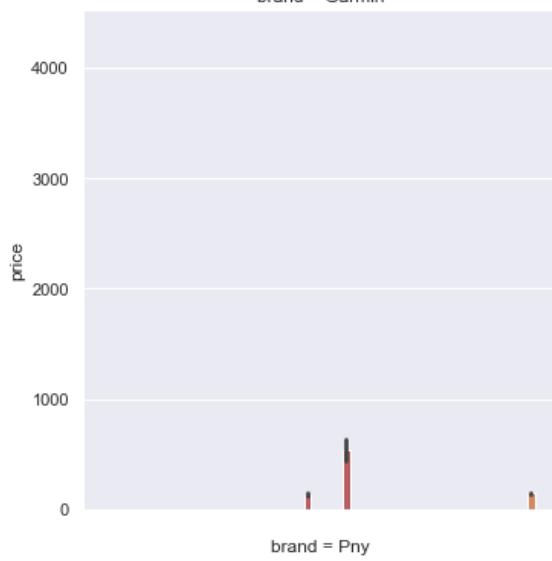
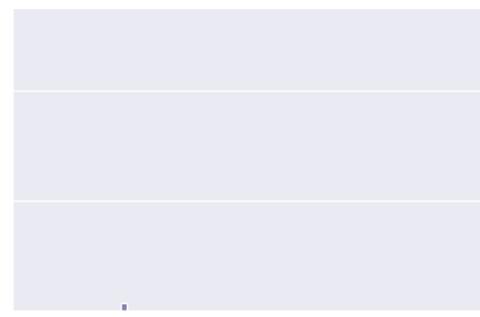
brand = Klipsch

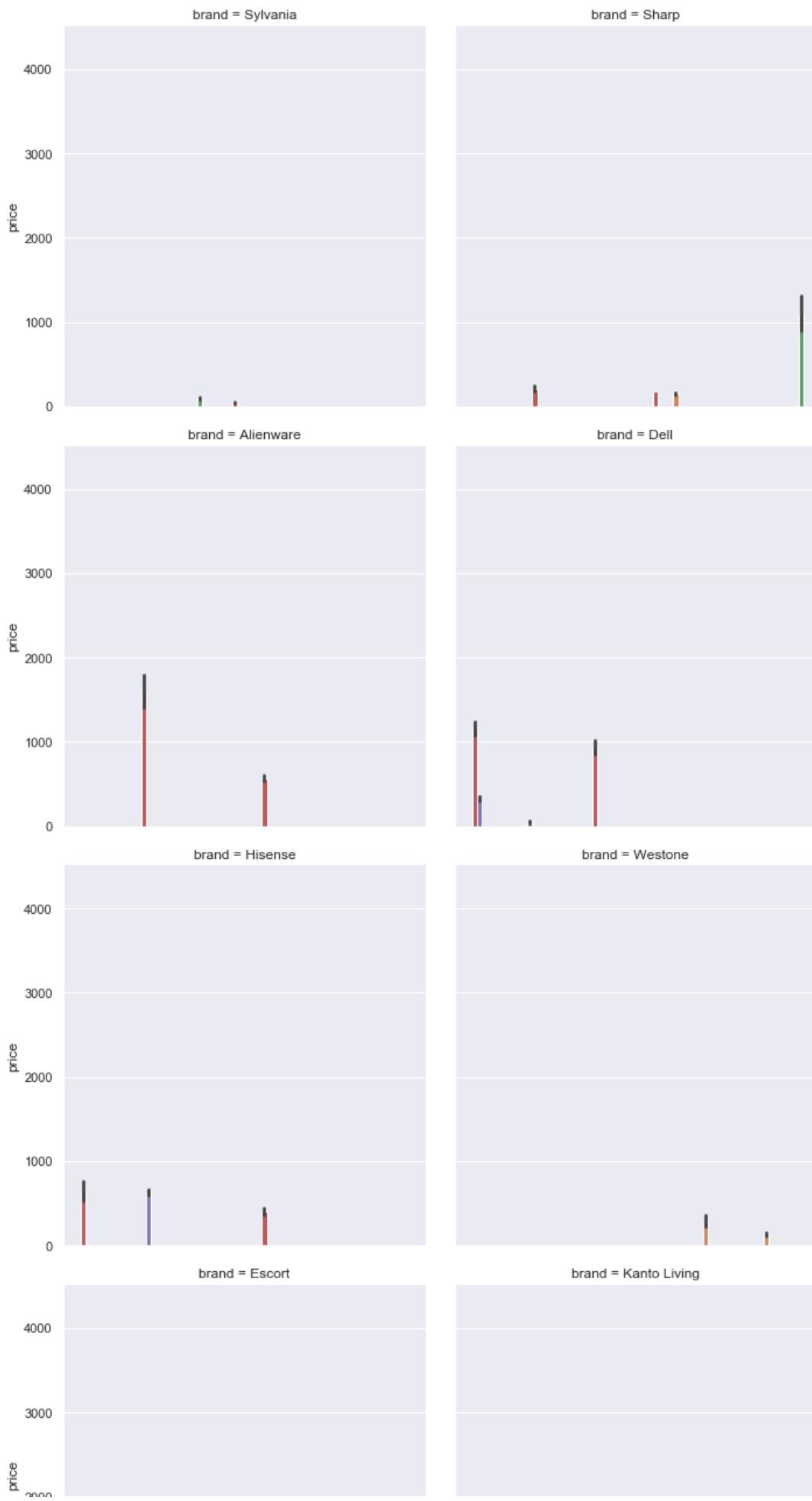


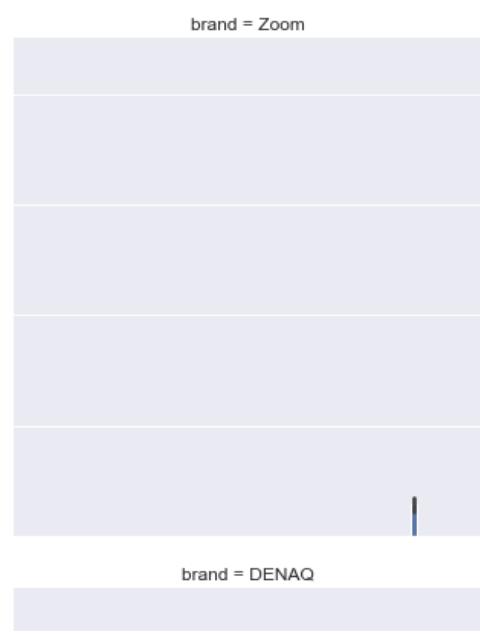
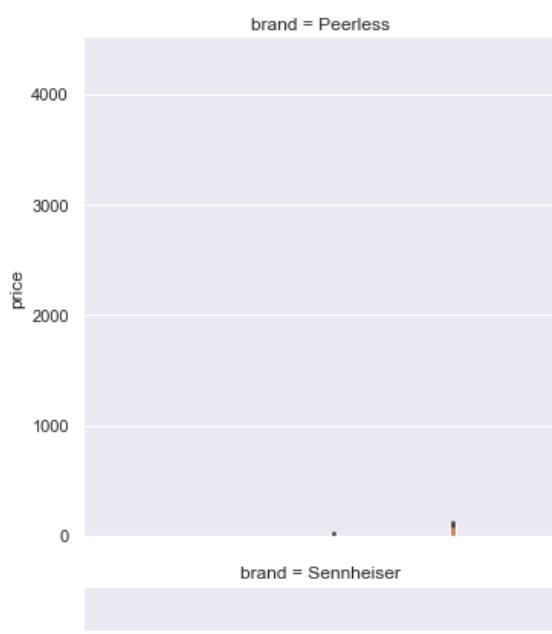
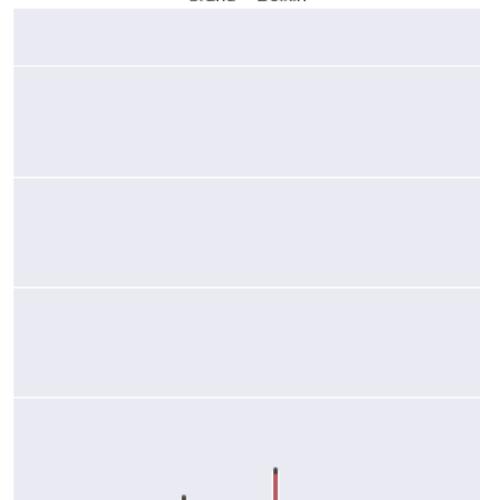
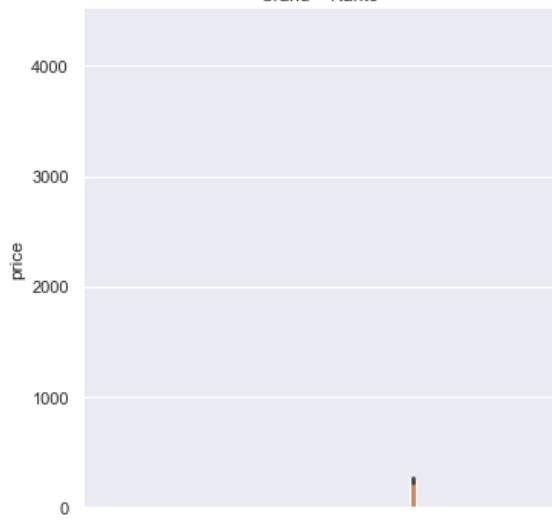
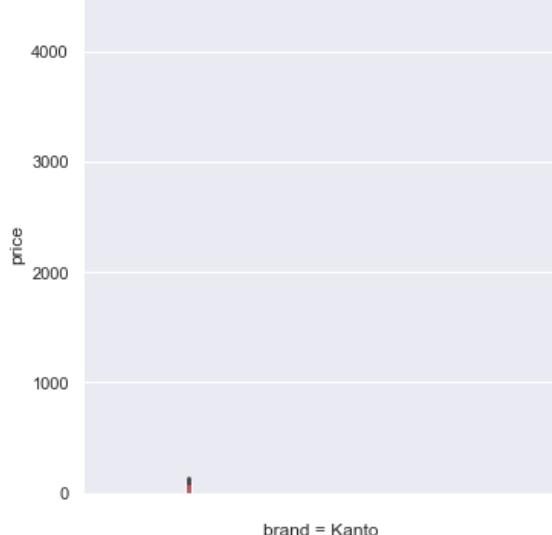
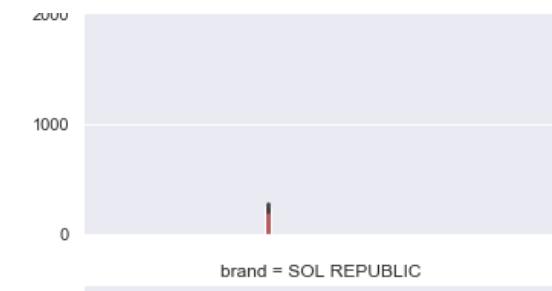


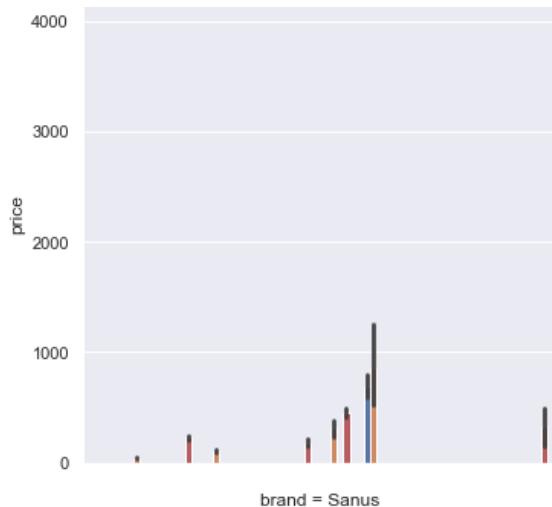




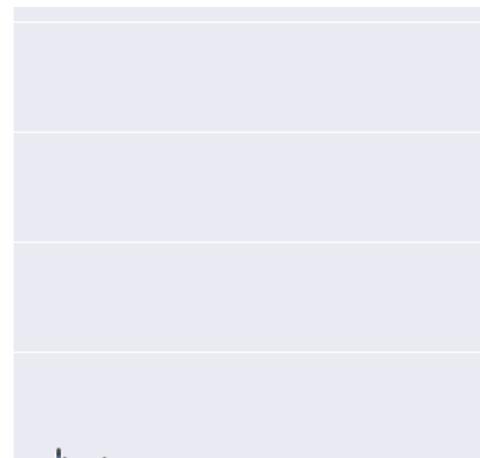




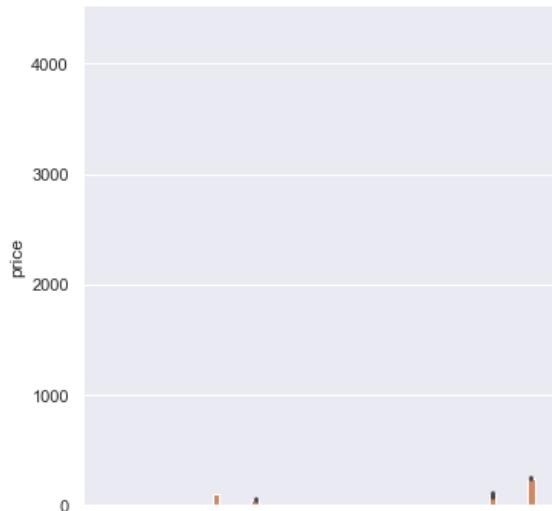




brand = Sanus



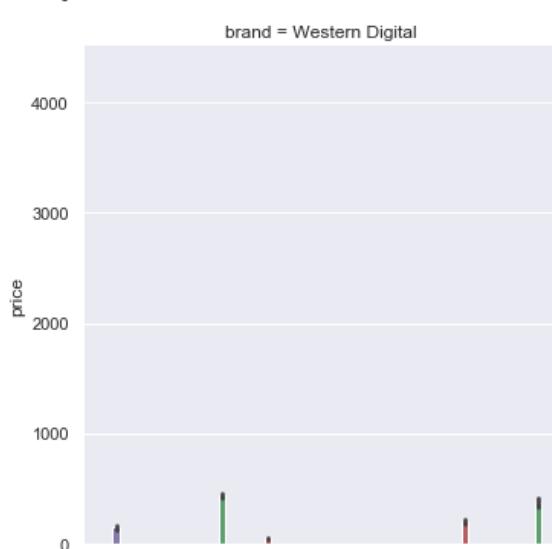
brand = Bose®



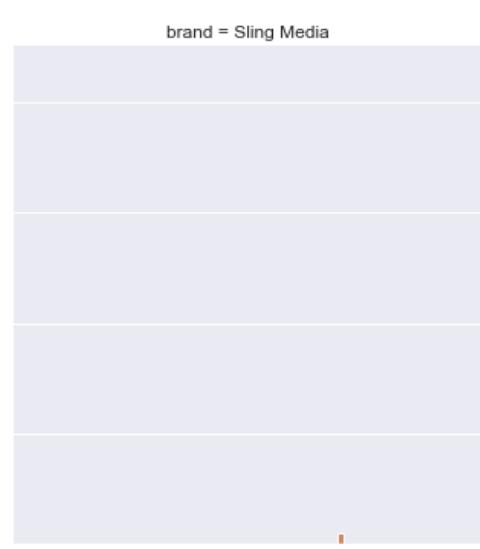
brand = Western Digital



brand = Sling Media

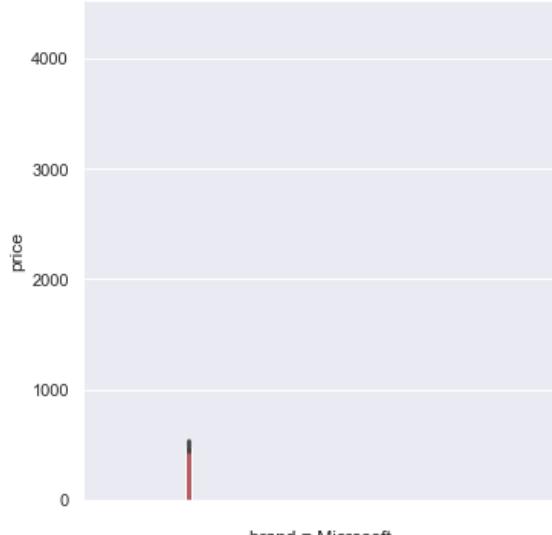
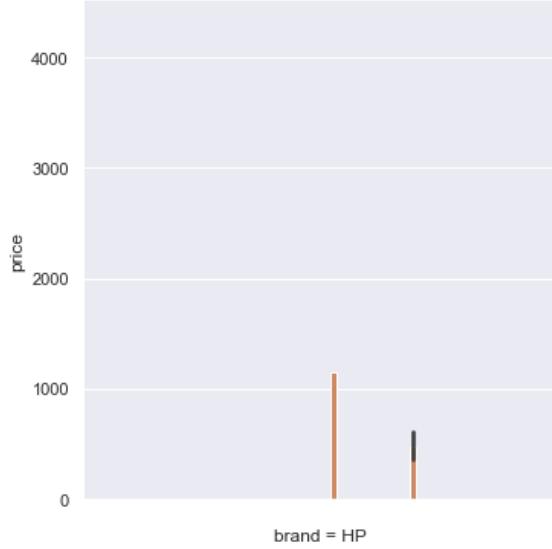
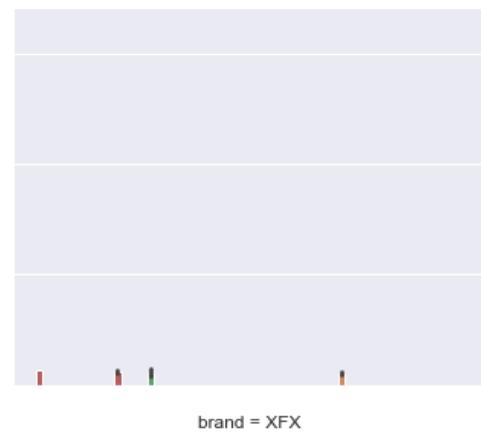
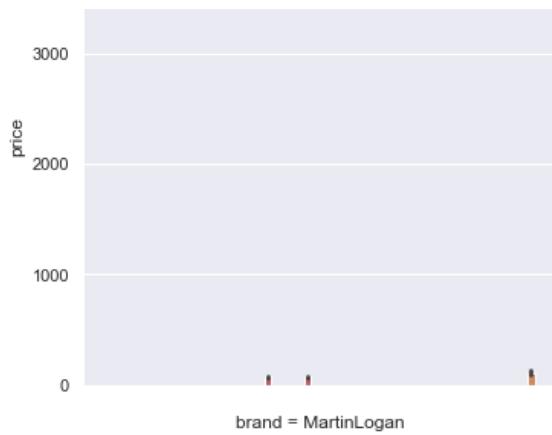


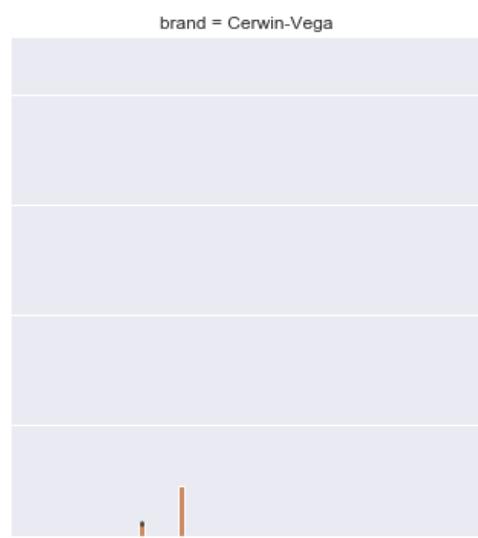
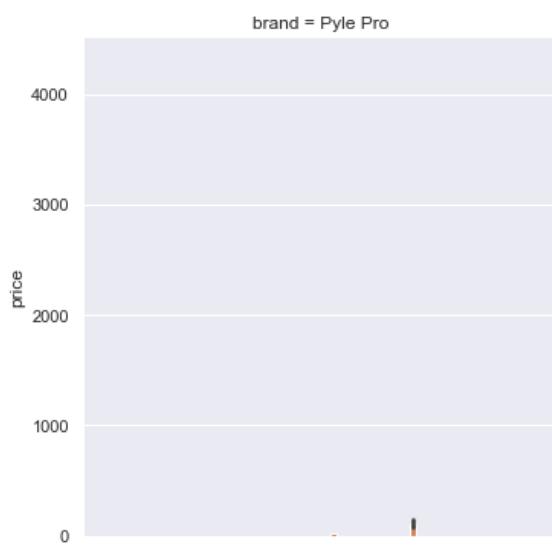
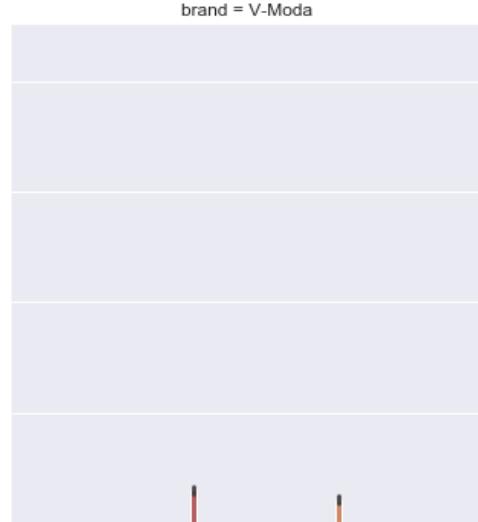
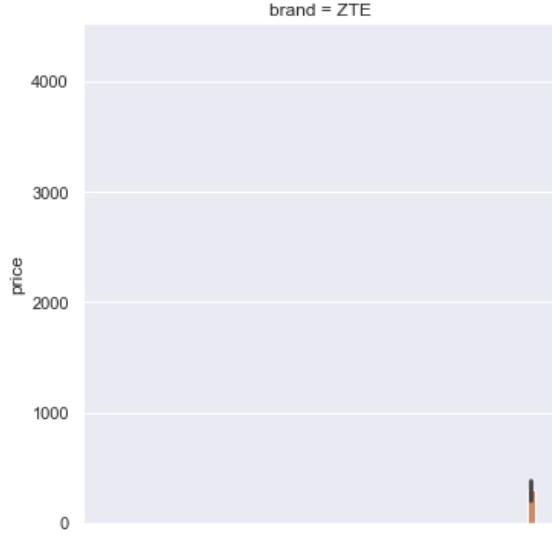
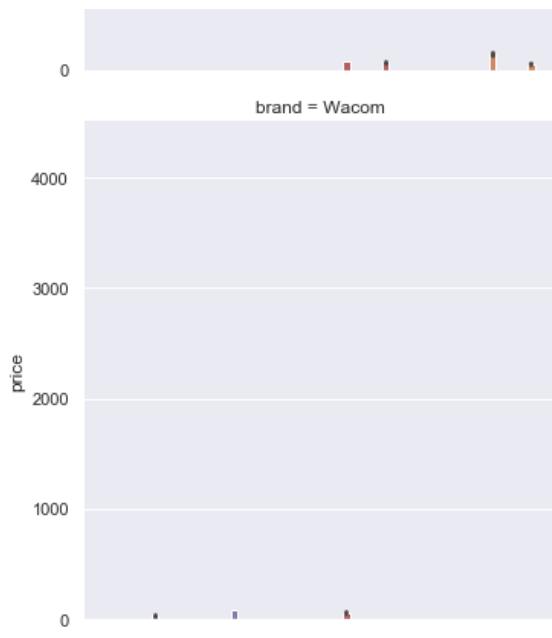
brand = Sima

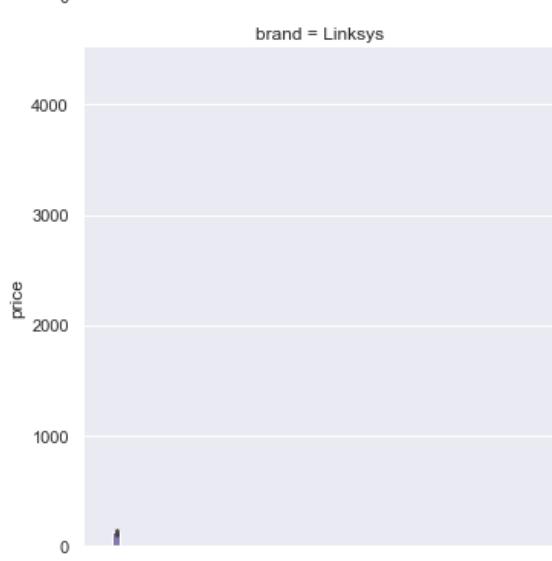
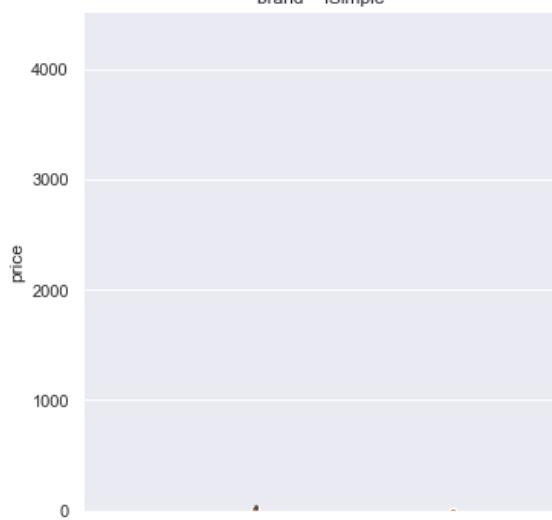
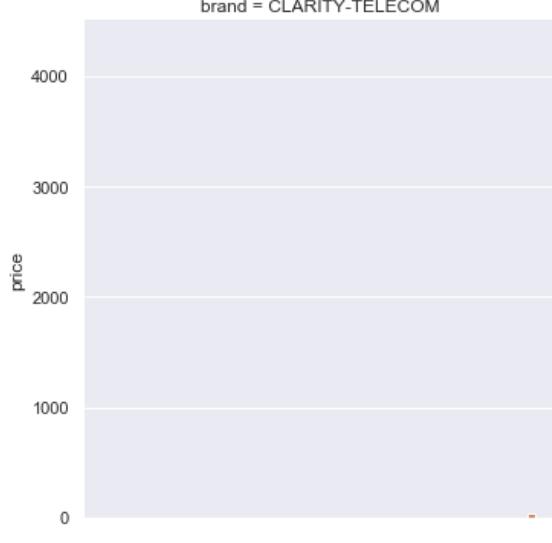
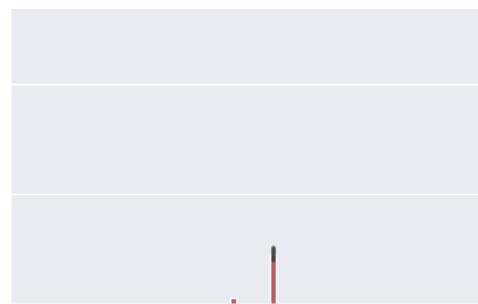


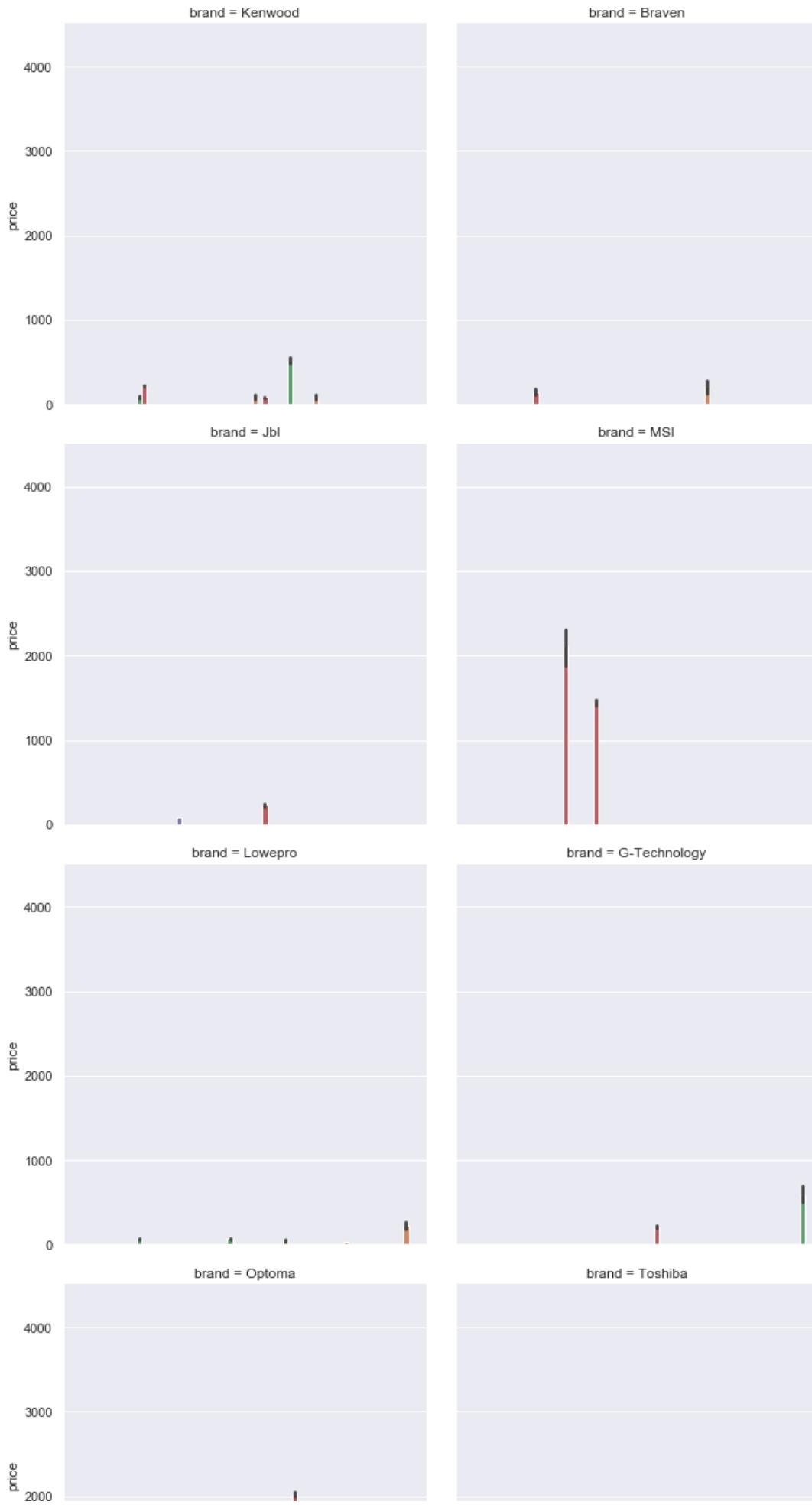
brand = Netgear

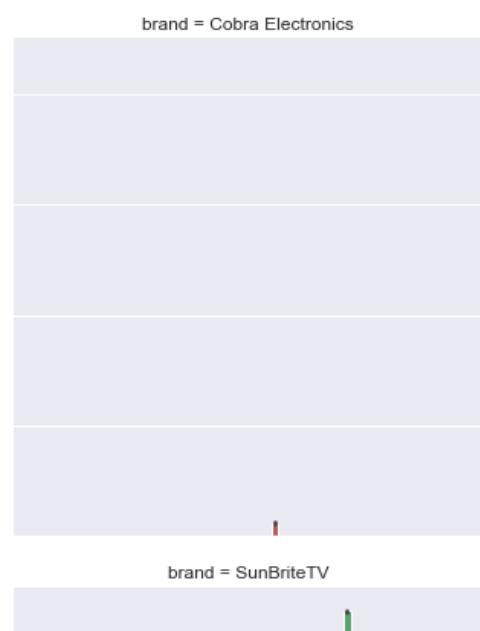
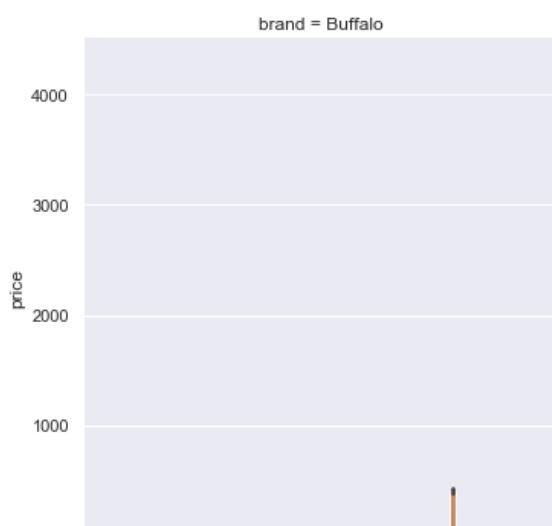
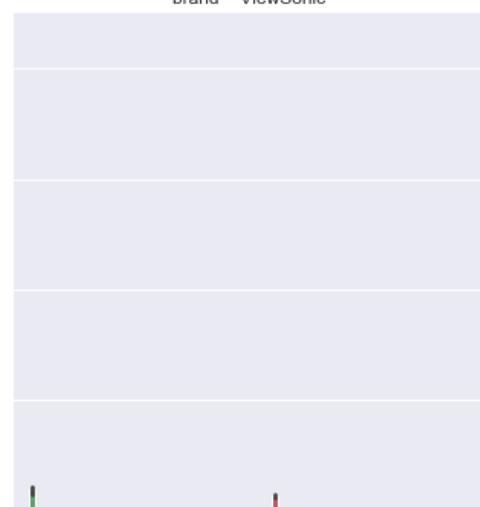
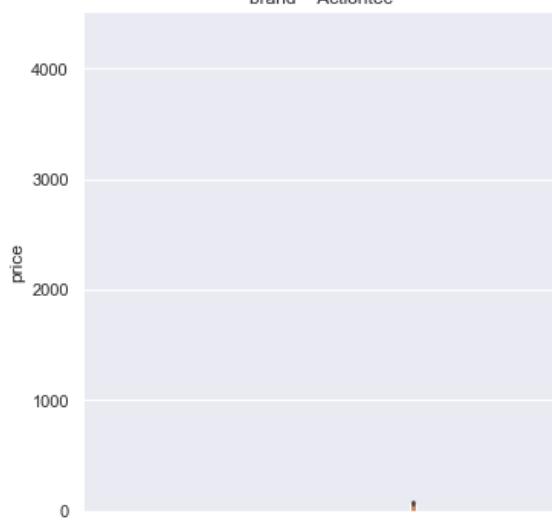
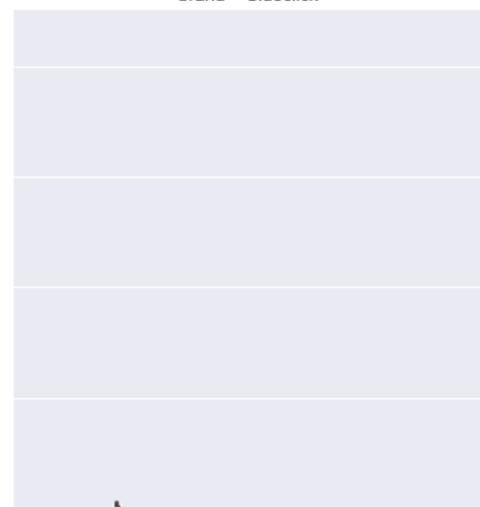
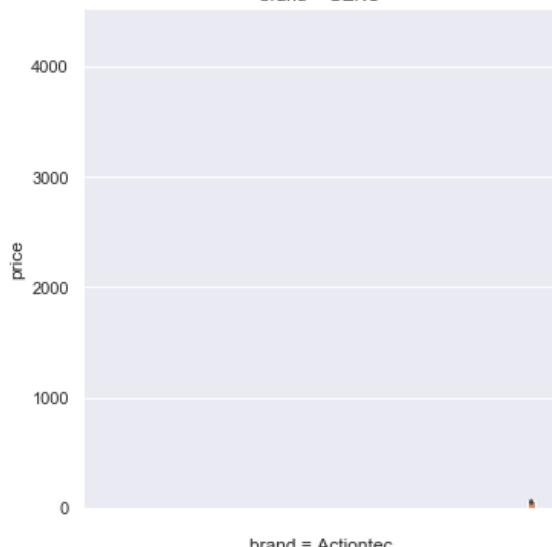
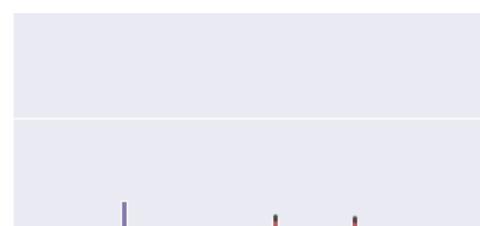


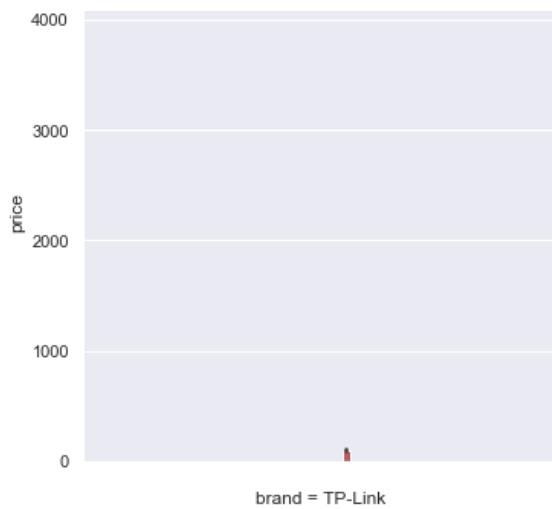




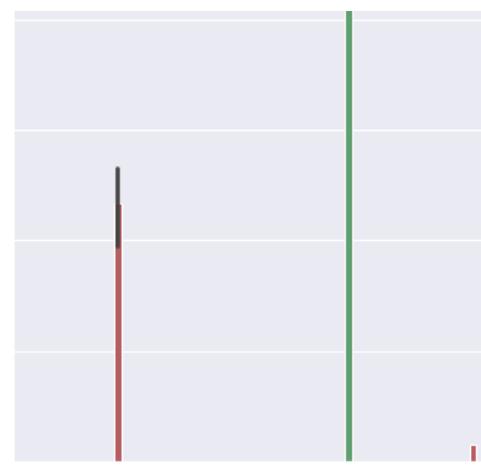




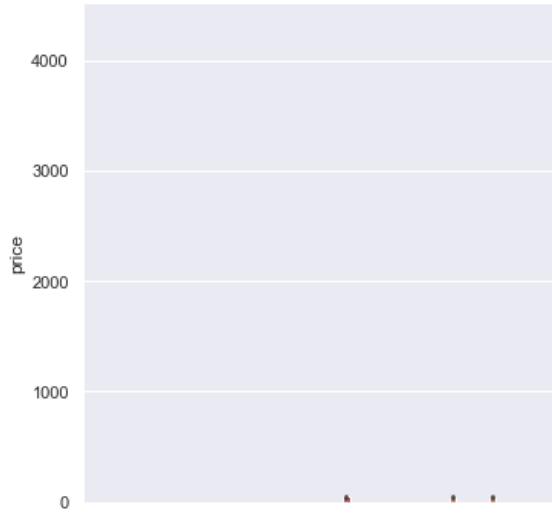




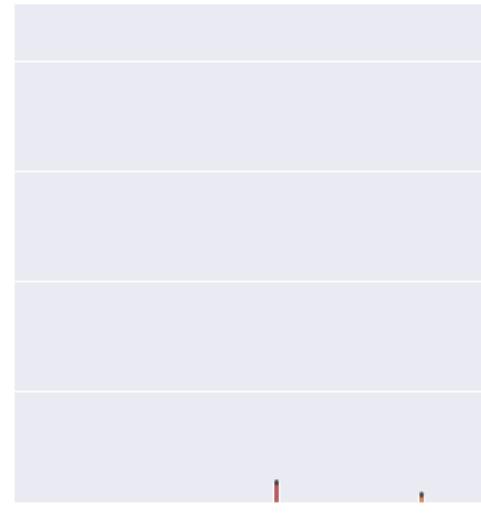
brand = TP-Link



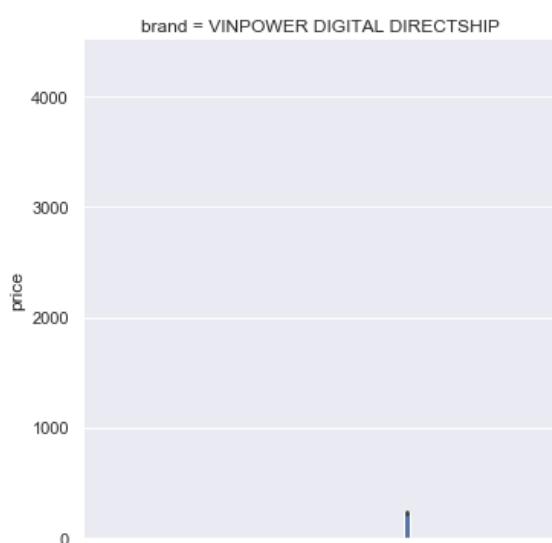
brand = Cooler Master



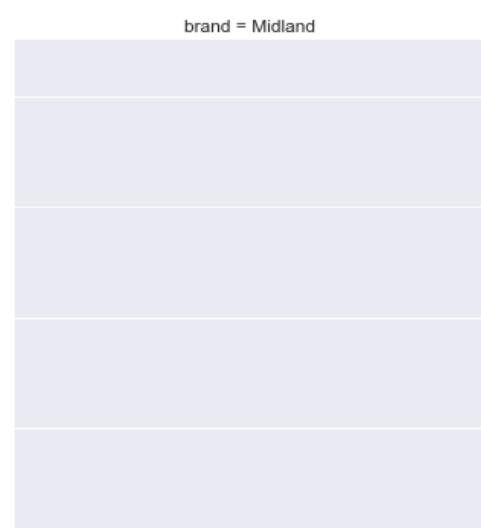
brand = VINPOWER DIGITAL DIRECTSHIP



brand = Midland

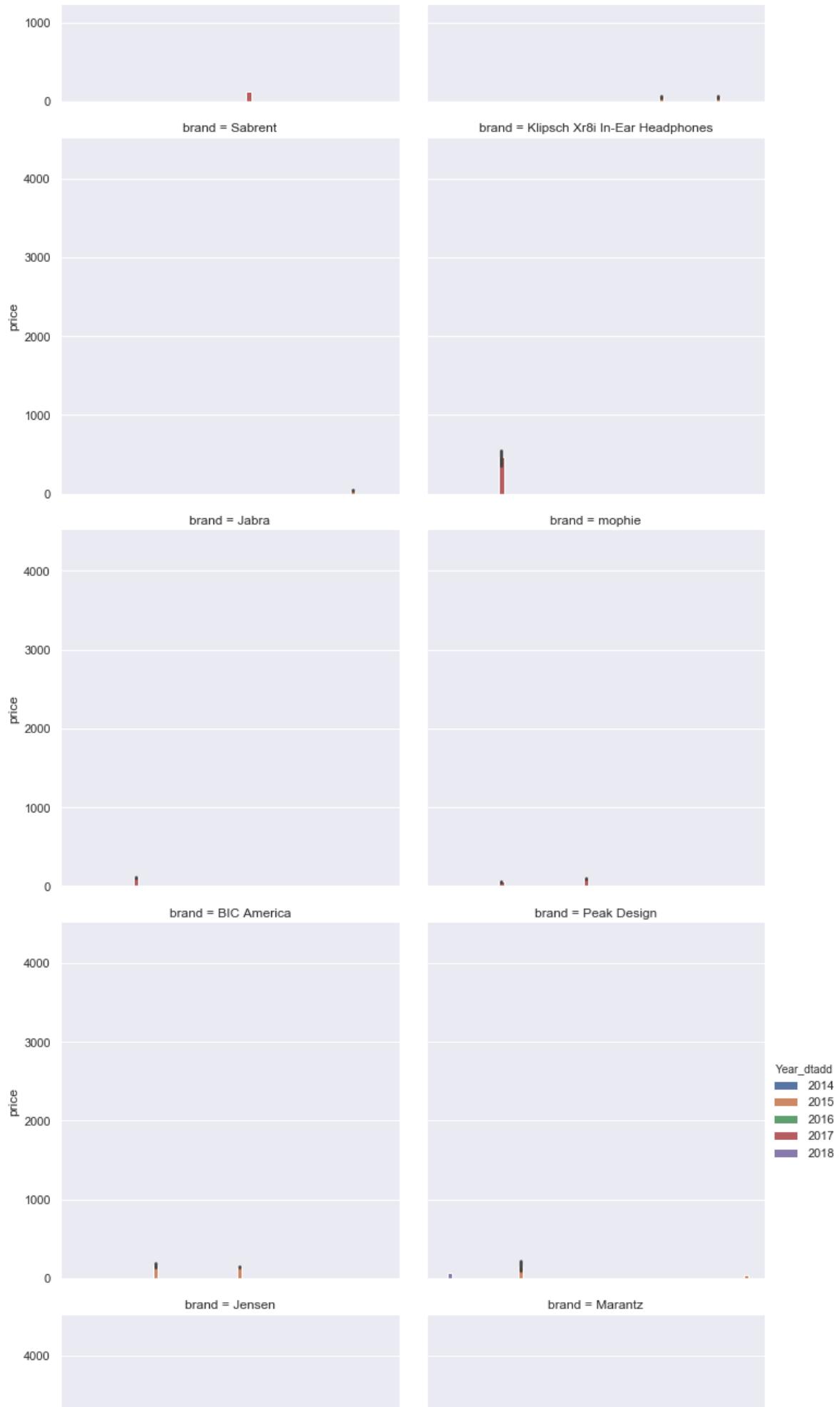


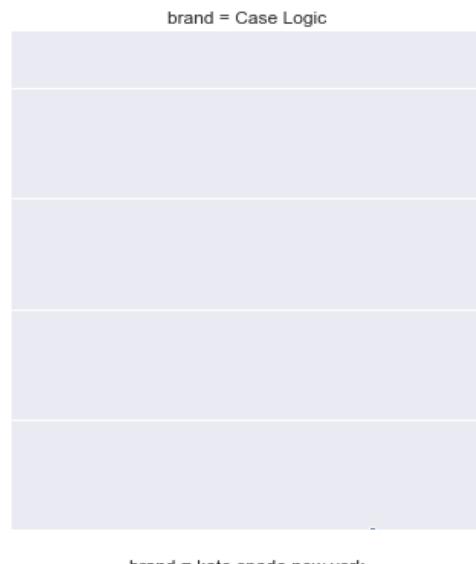
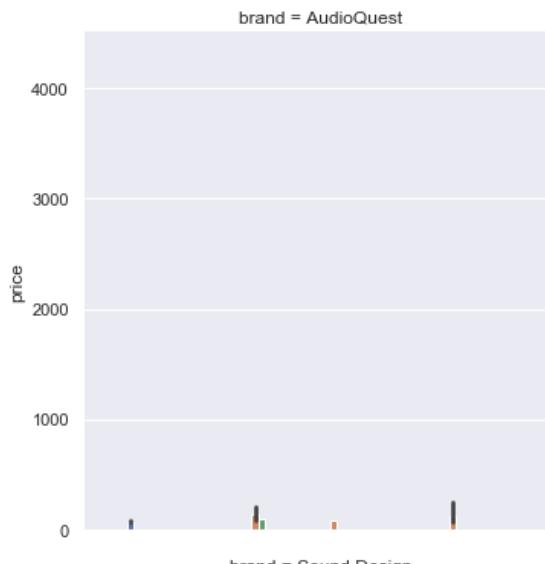
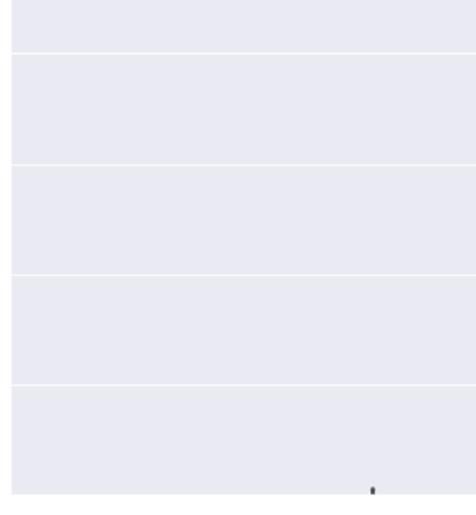
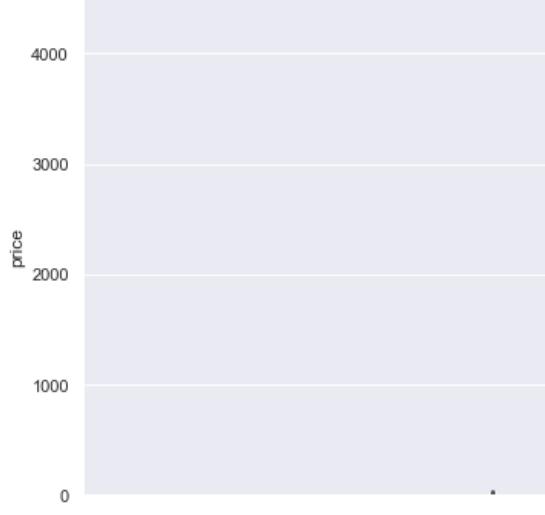
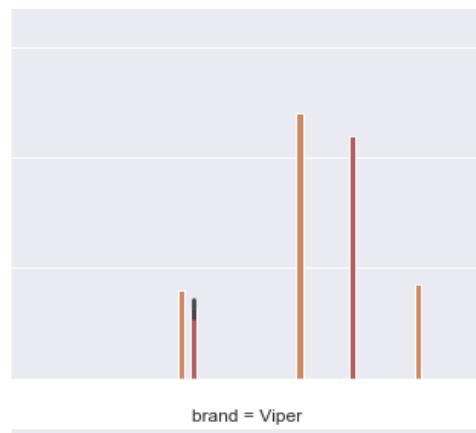
brand = Visidec

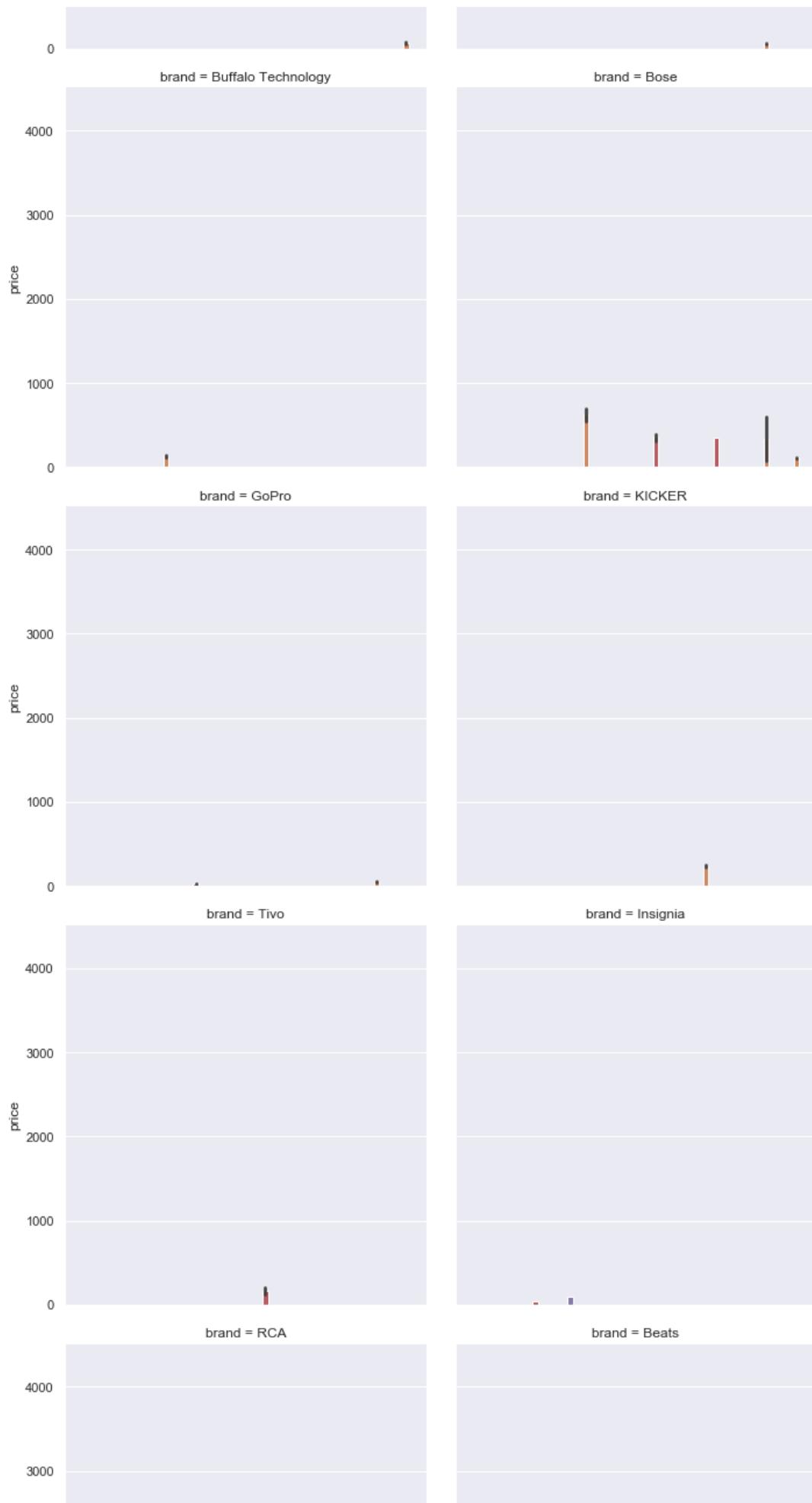


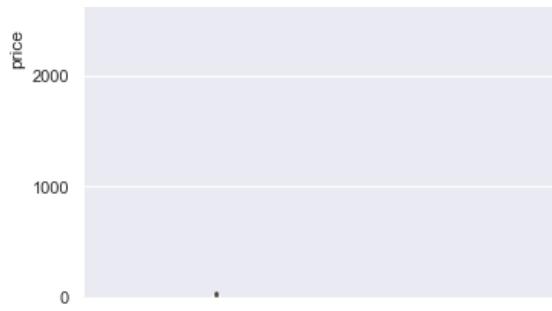
brand = Bower







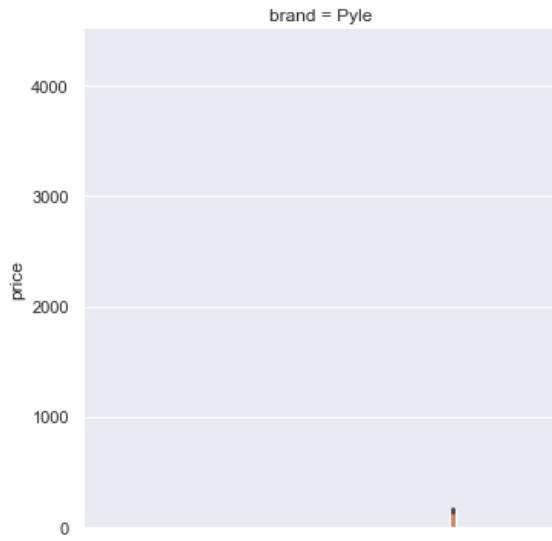




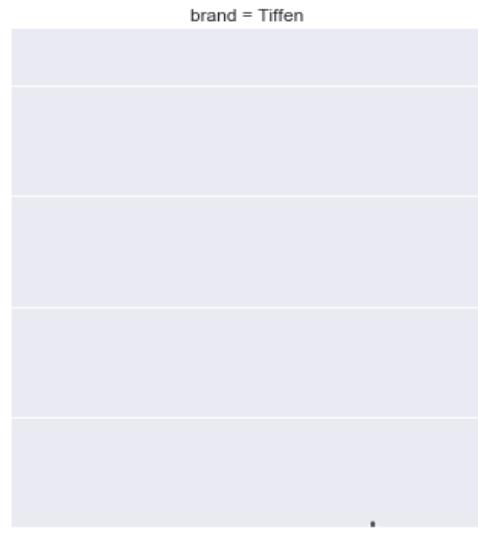
brand = Pyle



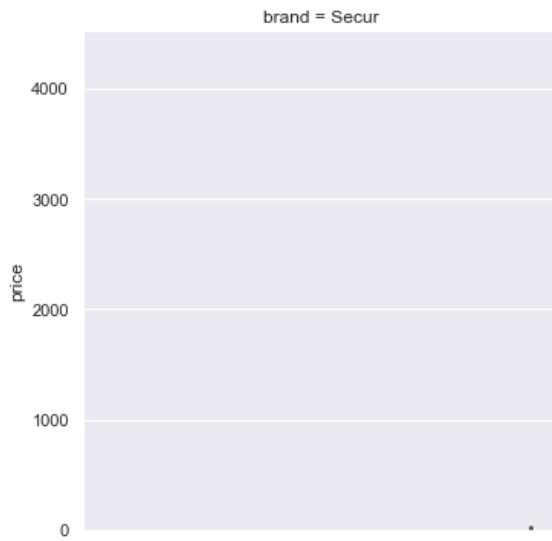
brand = Tiffen



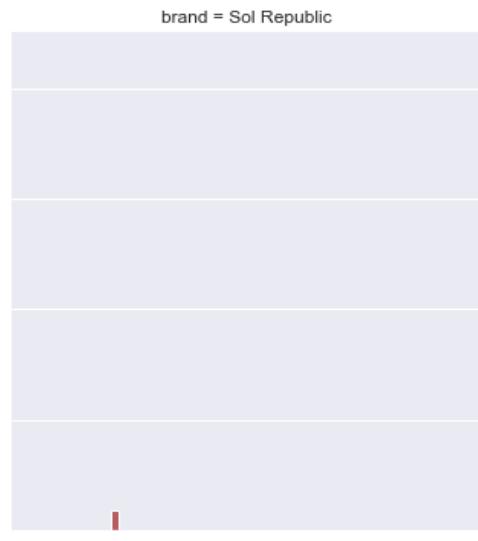
brand = Secur



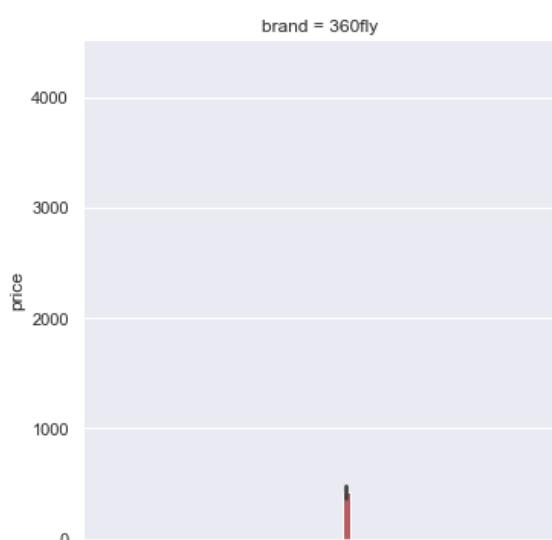
brand = Sol Republic



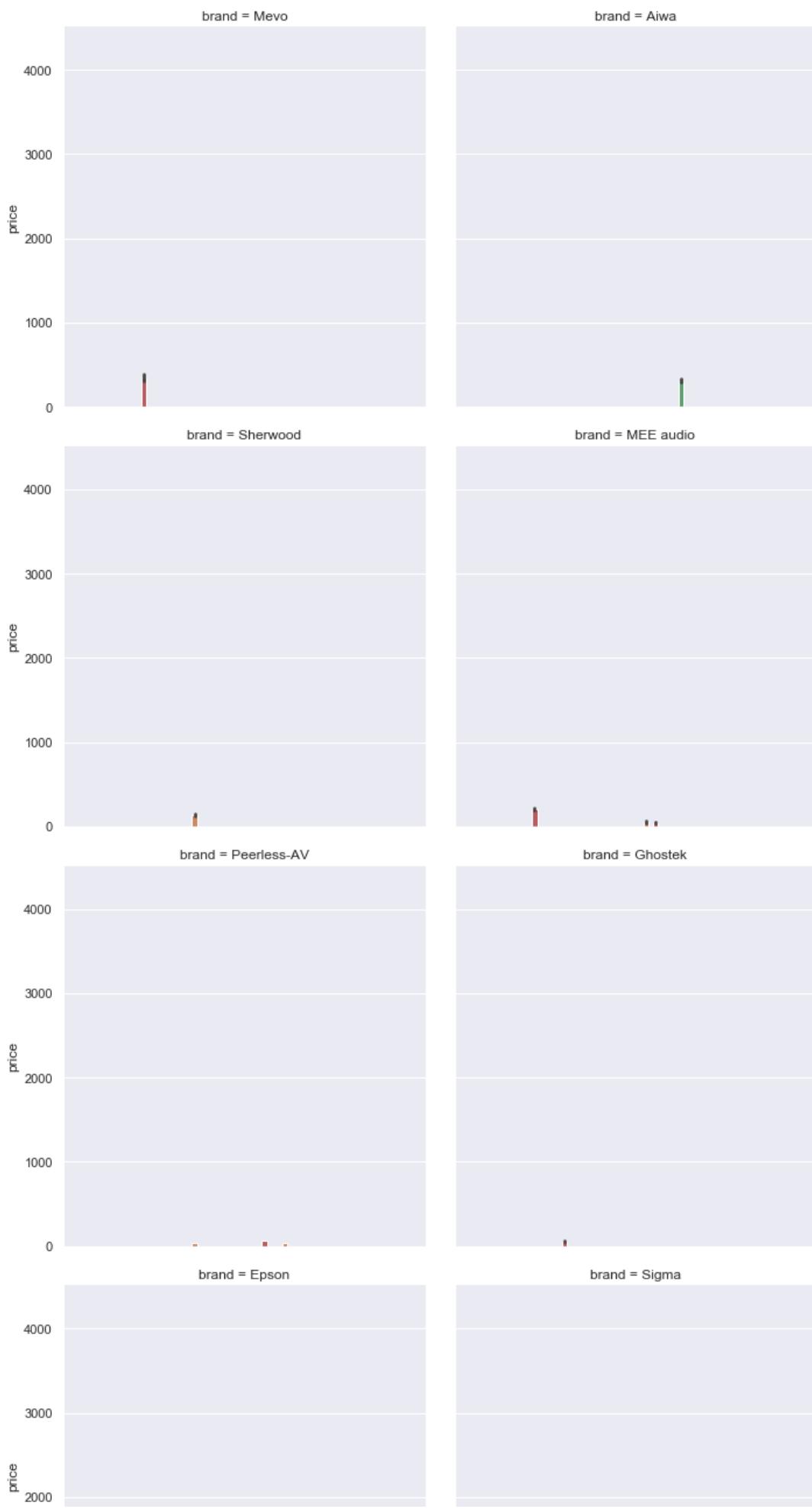
brand = 360fly

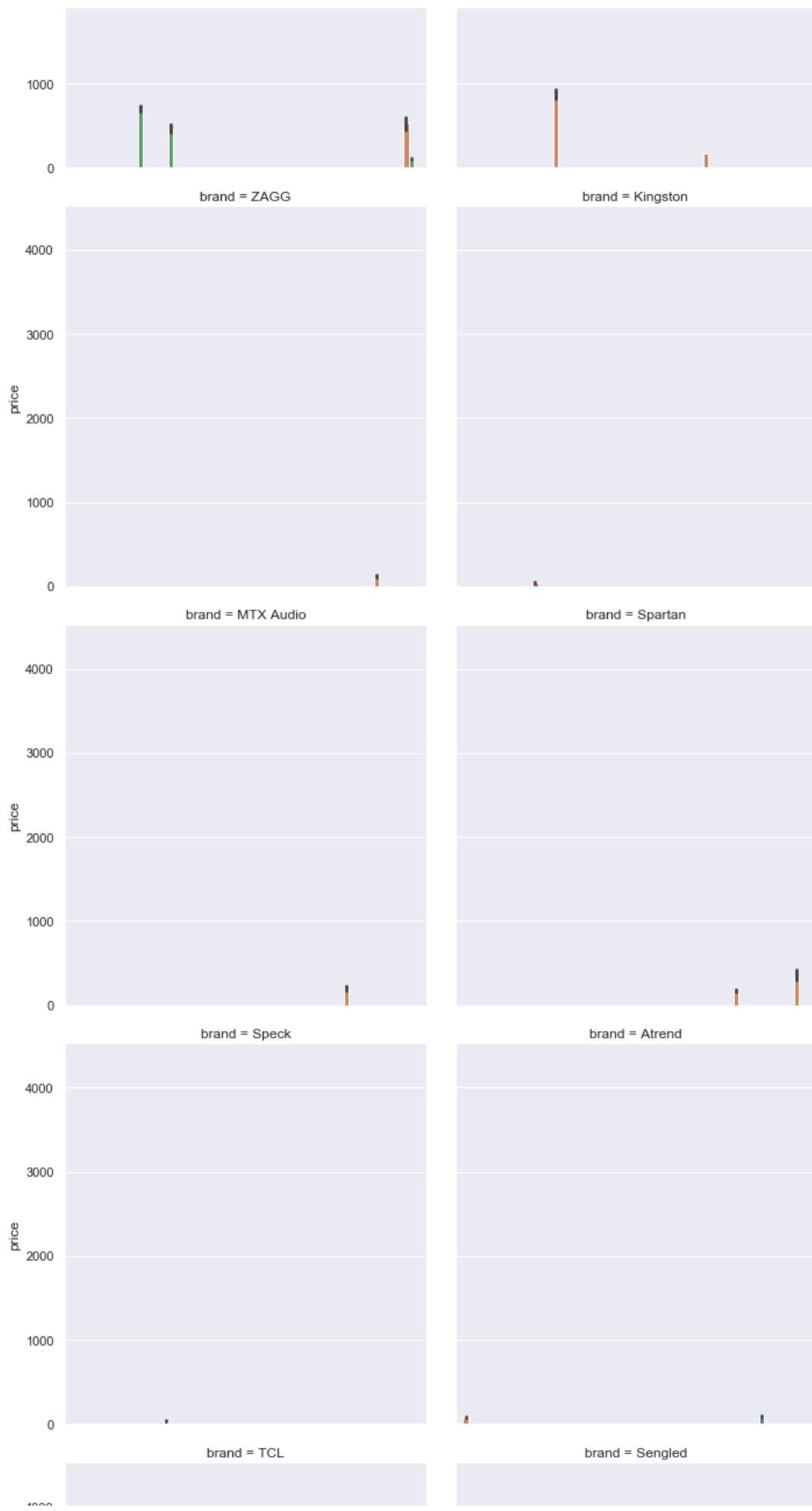


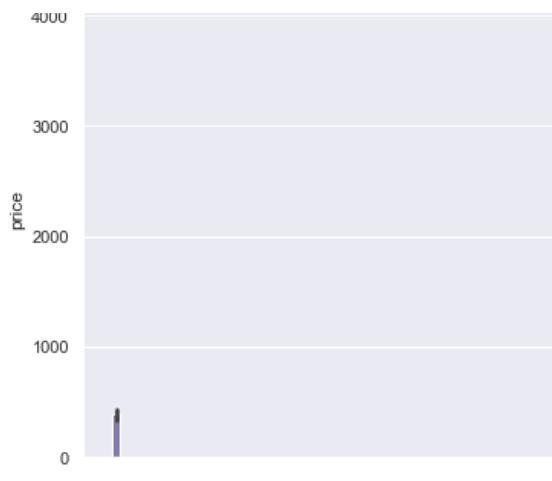
brand = Steelseries



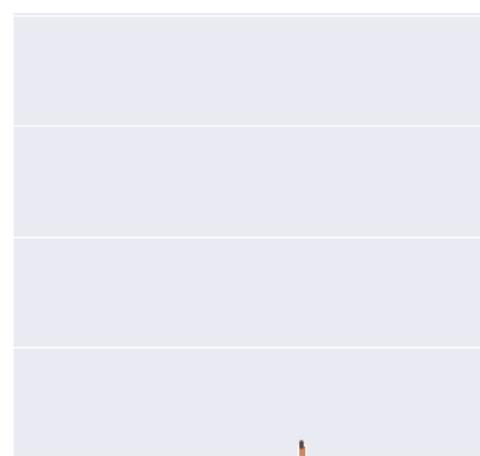
0



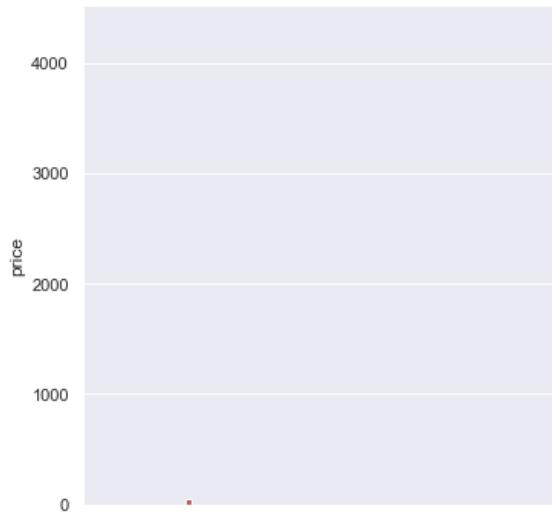




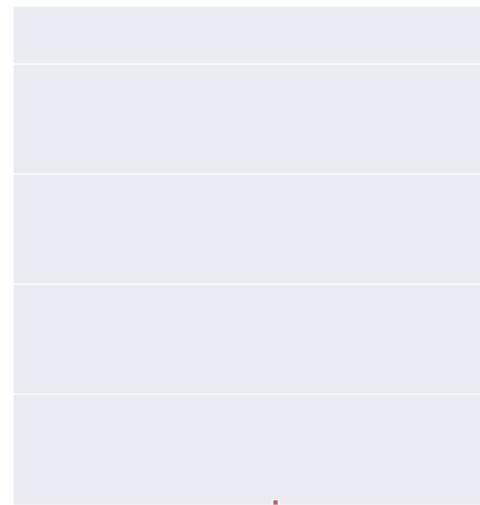
brand = Memorex



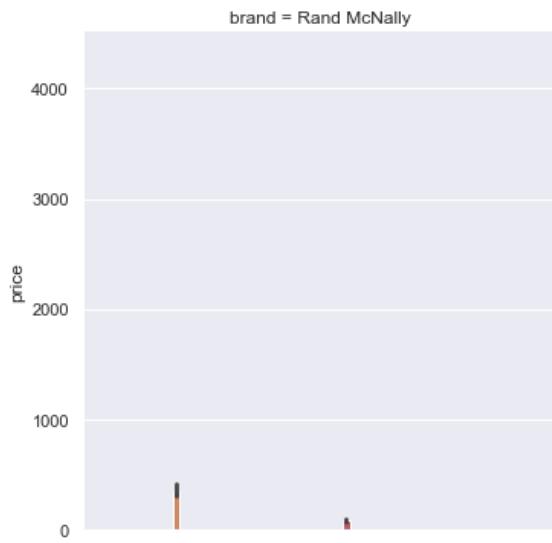
brand = Antec



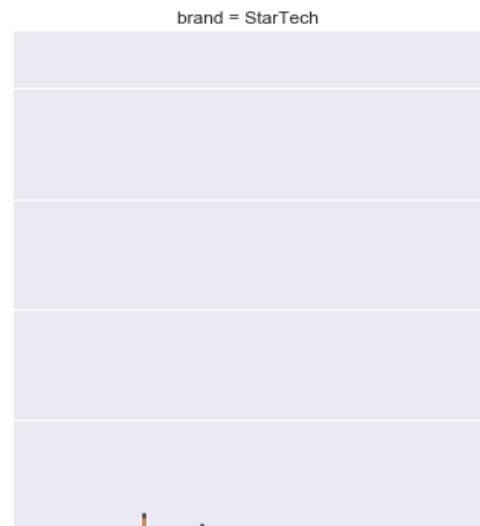
brand = Rand McNally



brand = StarTech

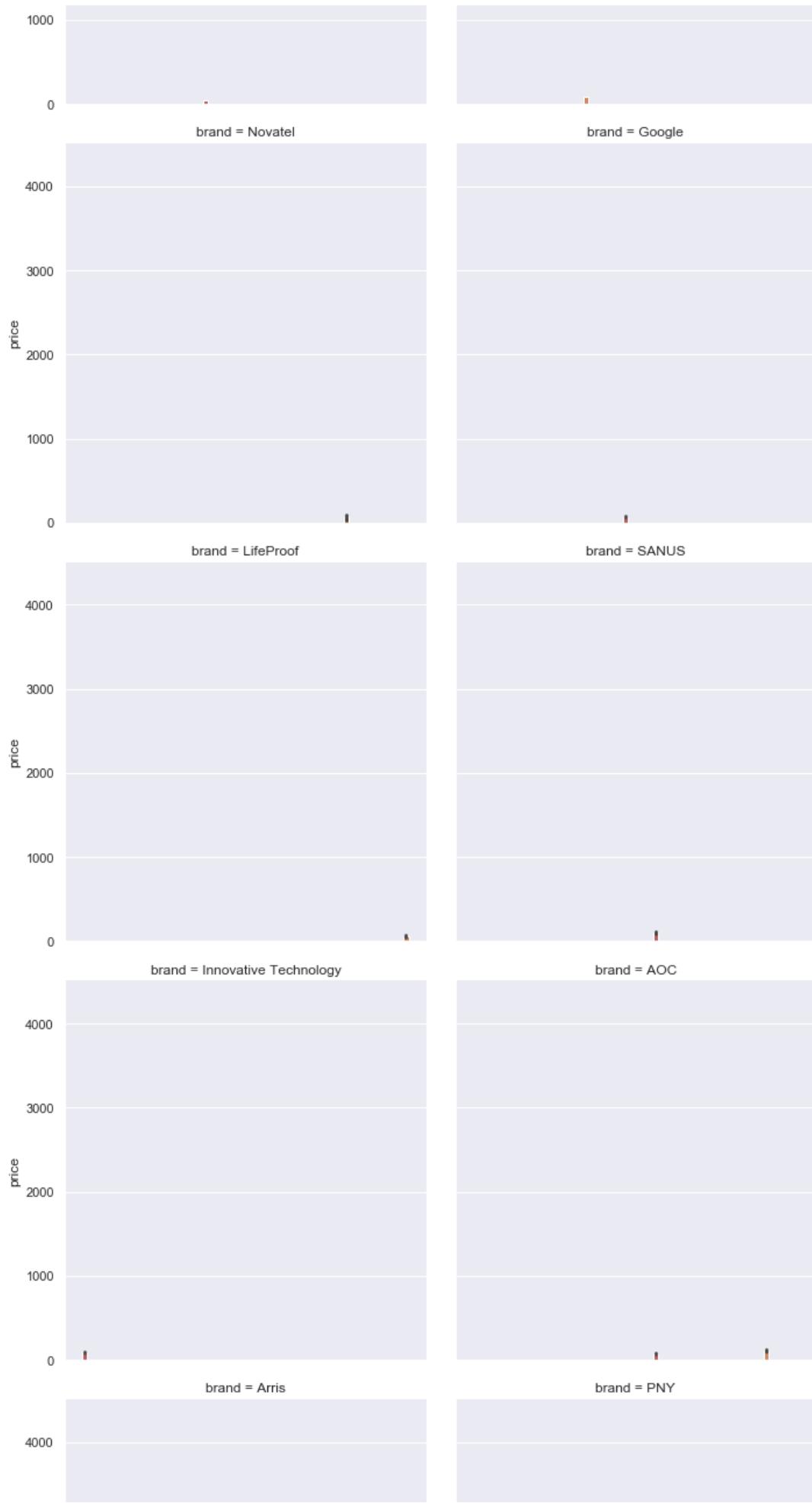


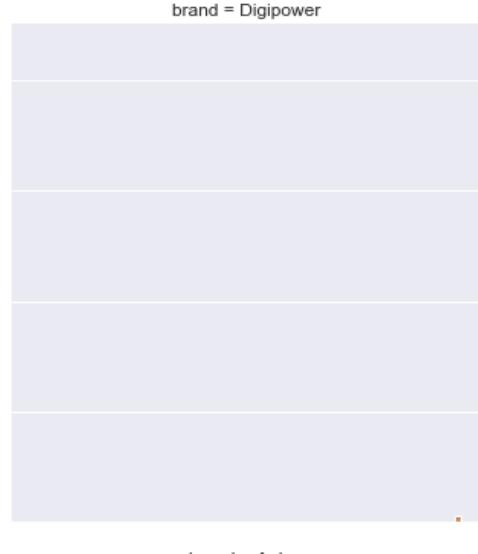
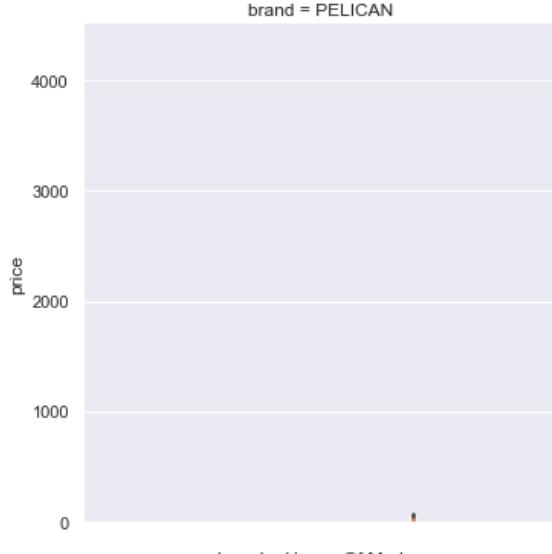
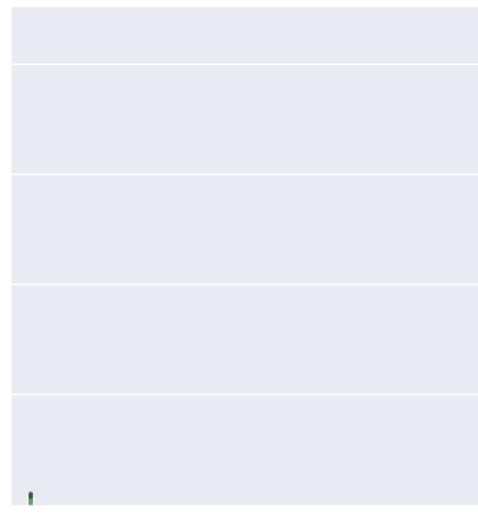
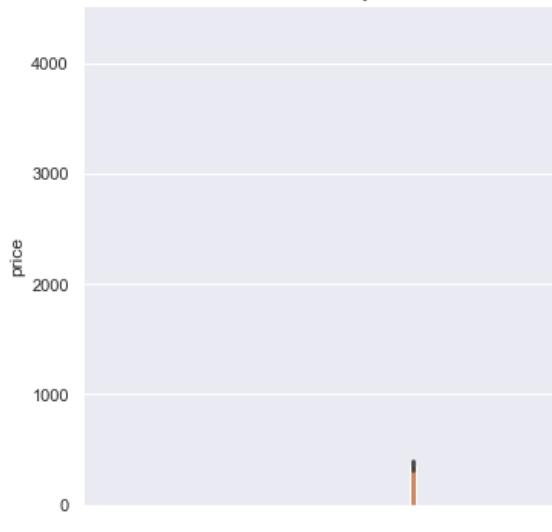
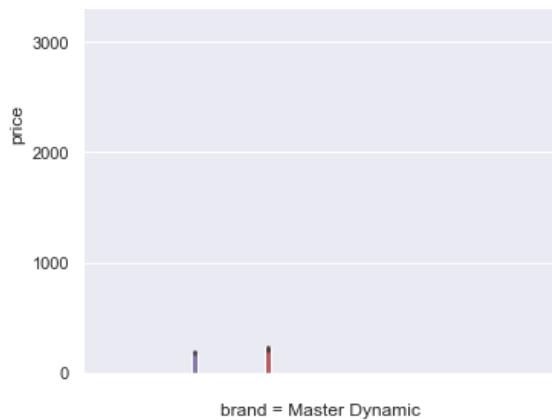
brand = Aftershokz

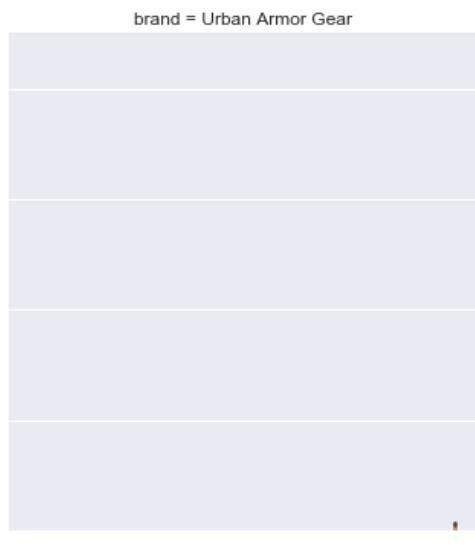
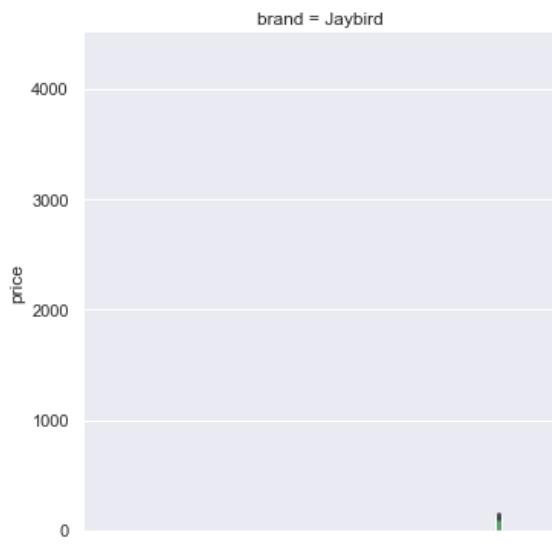
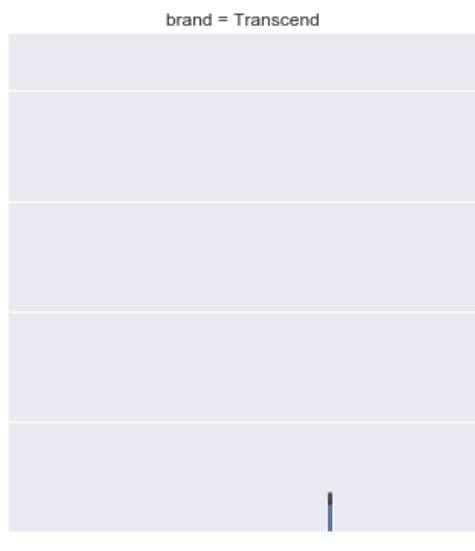
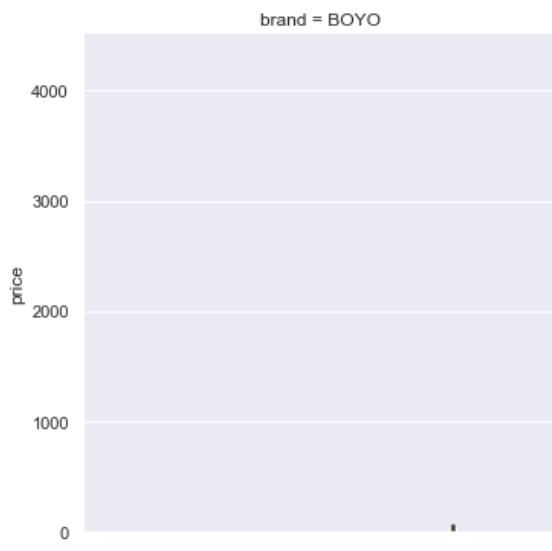
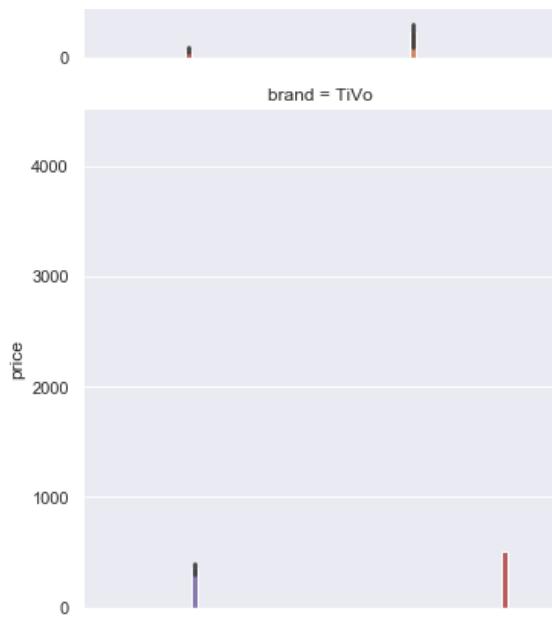


brand = MTX











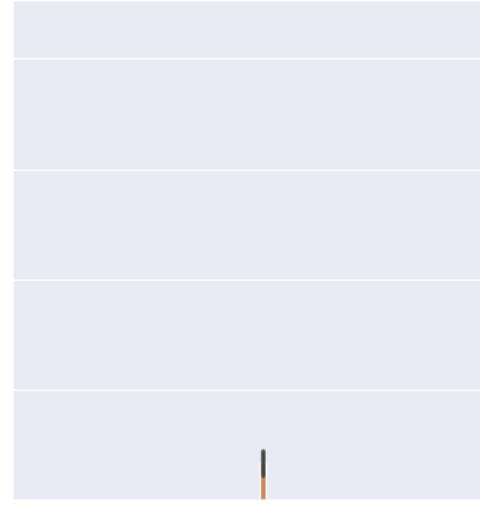
brand = Audioquest



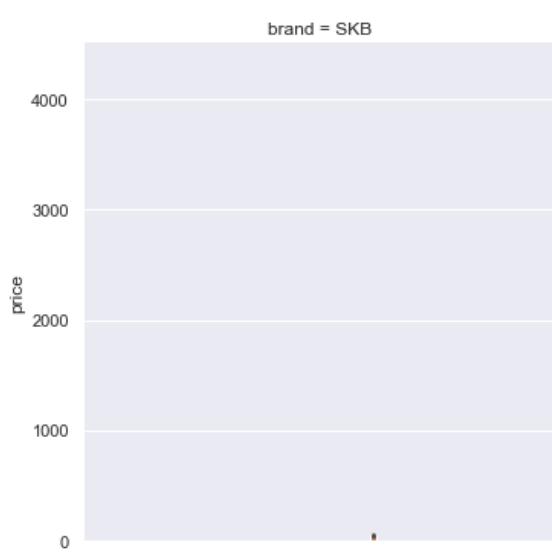
brand = SpeakerCraft



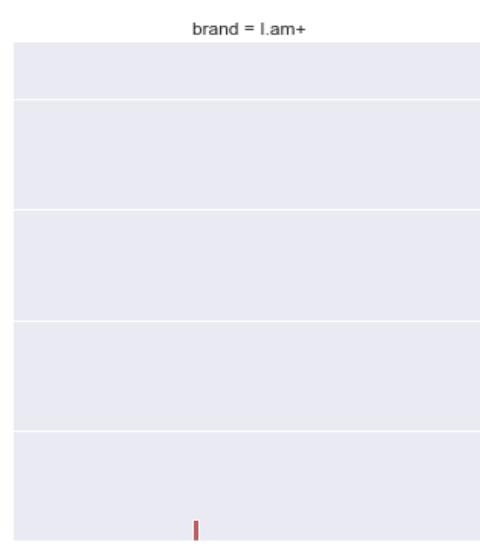
brand = SKB



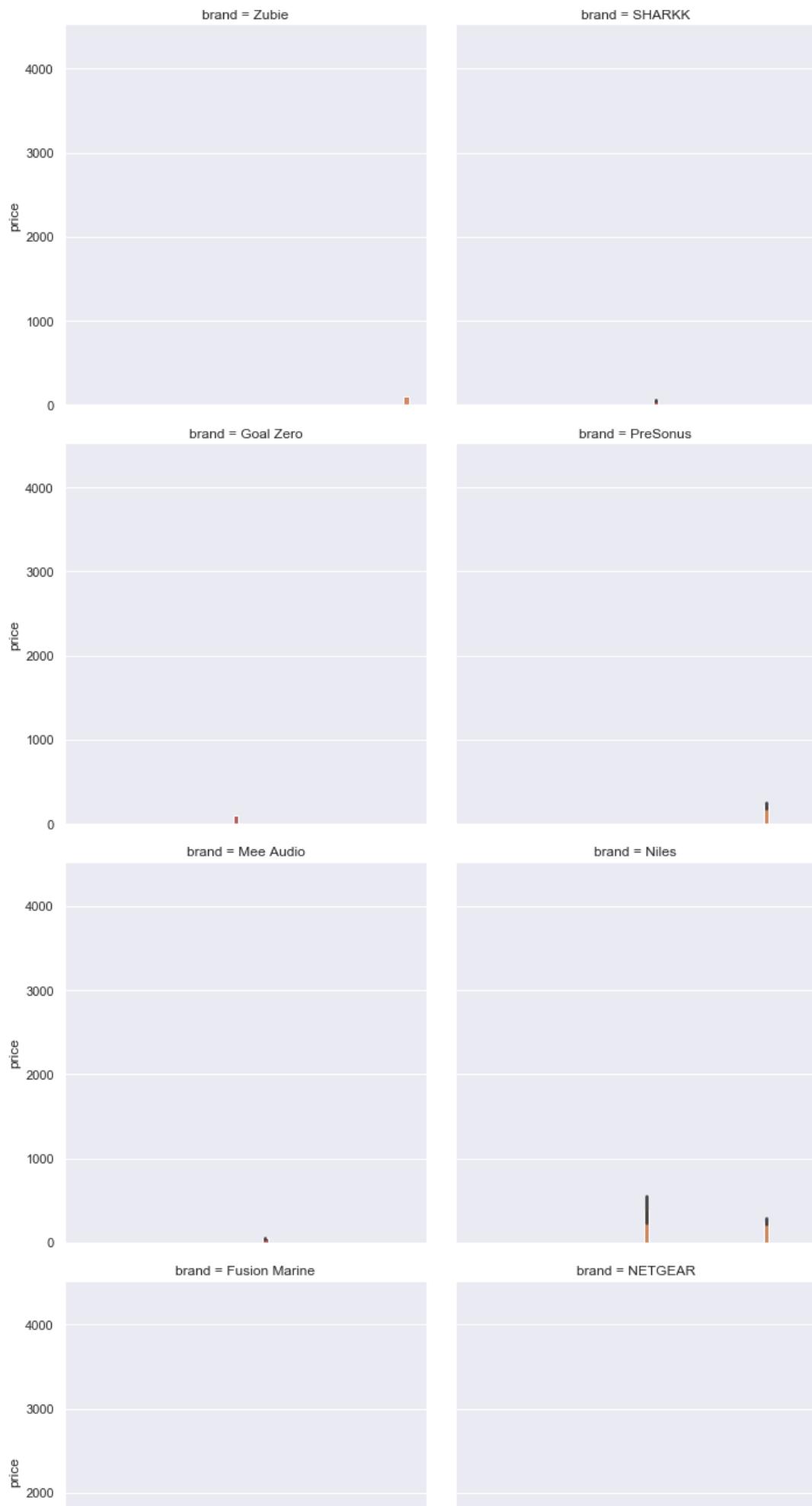
brand = I.am+

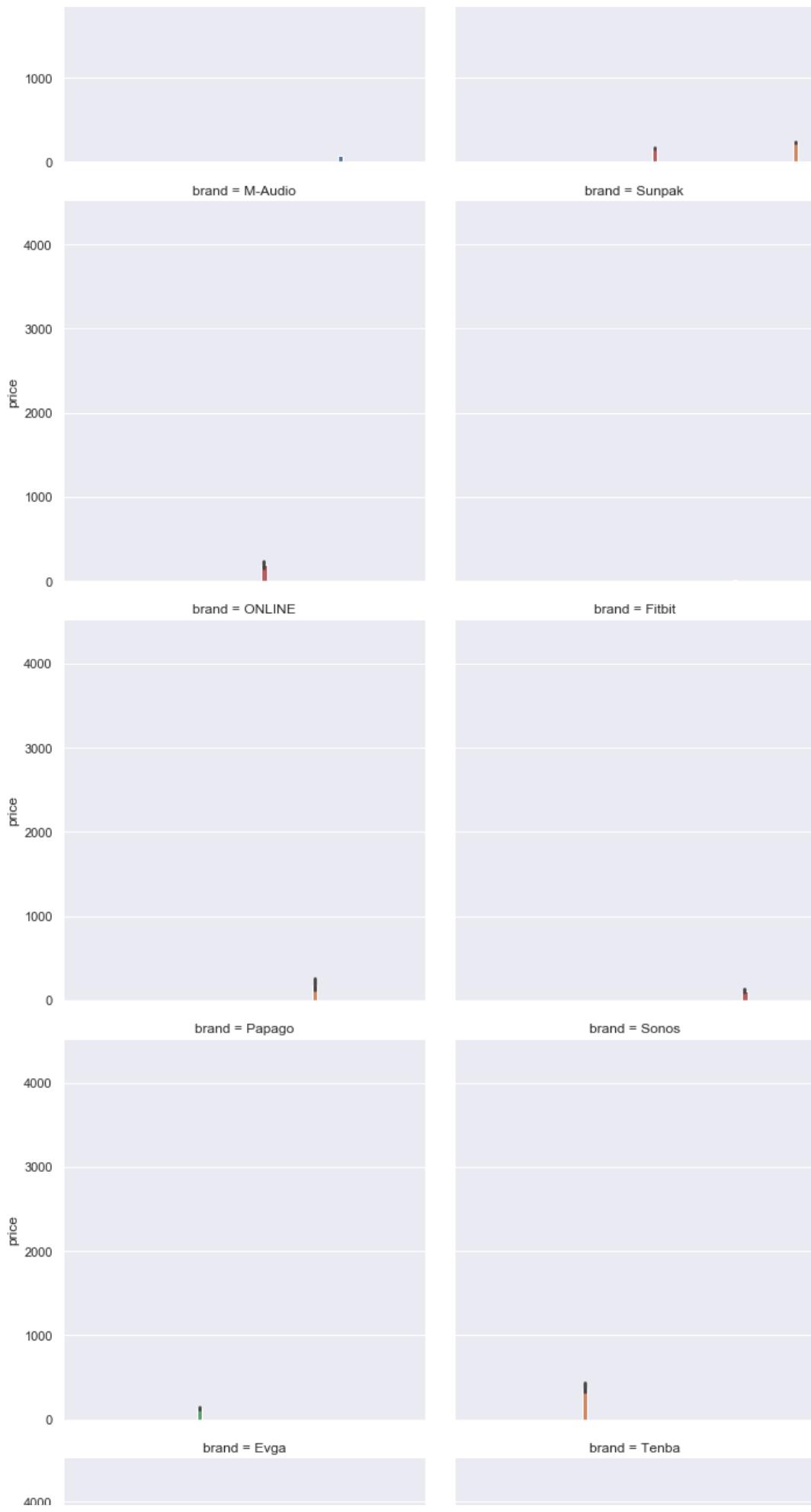


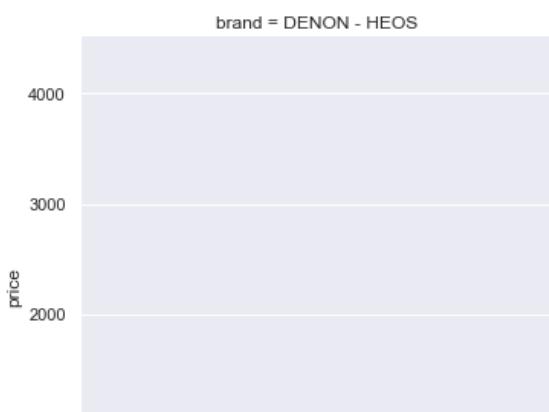
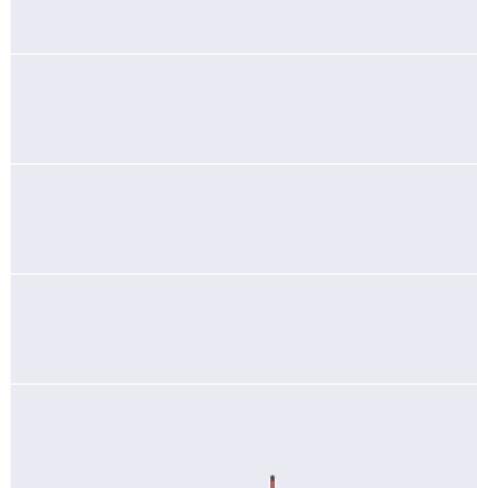
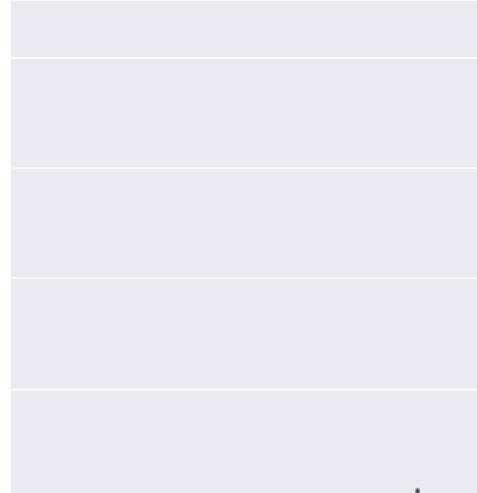
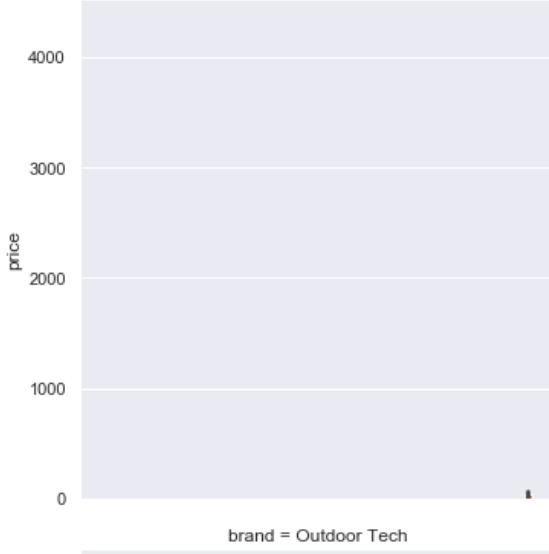
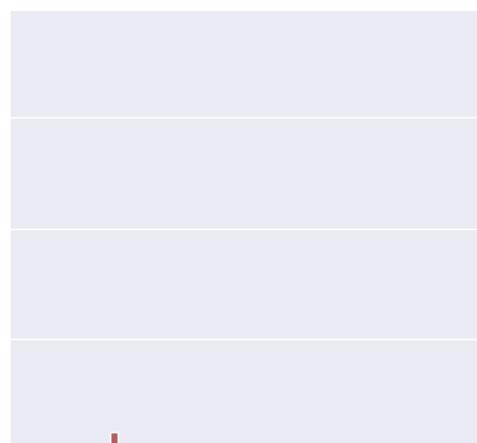
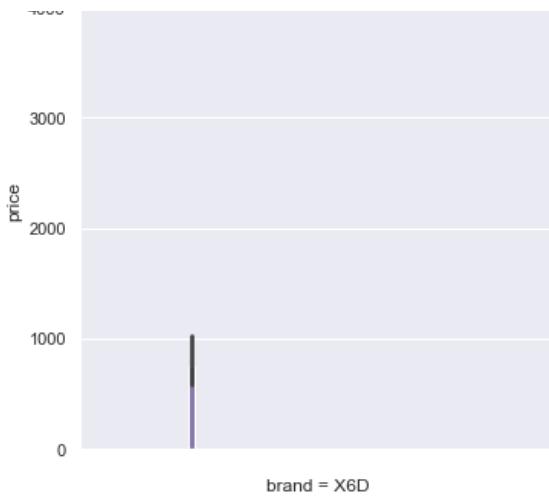
brand = Planet Audio

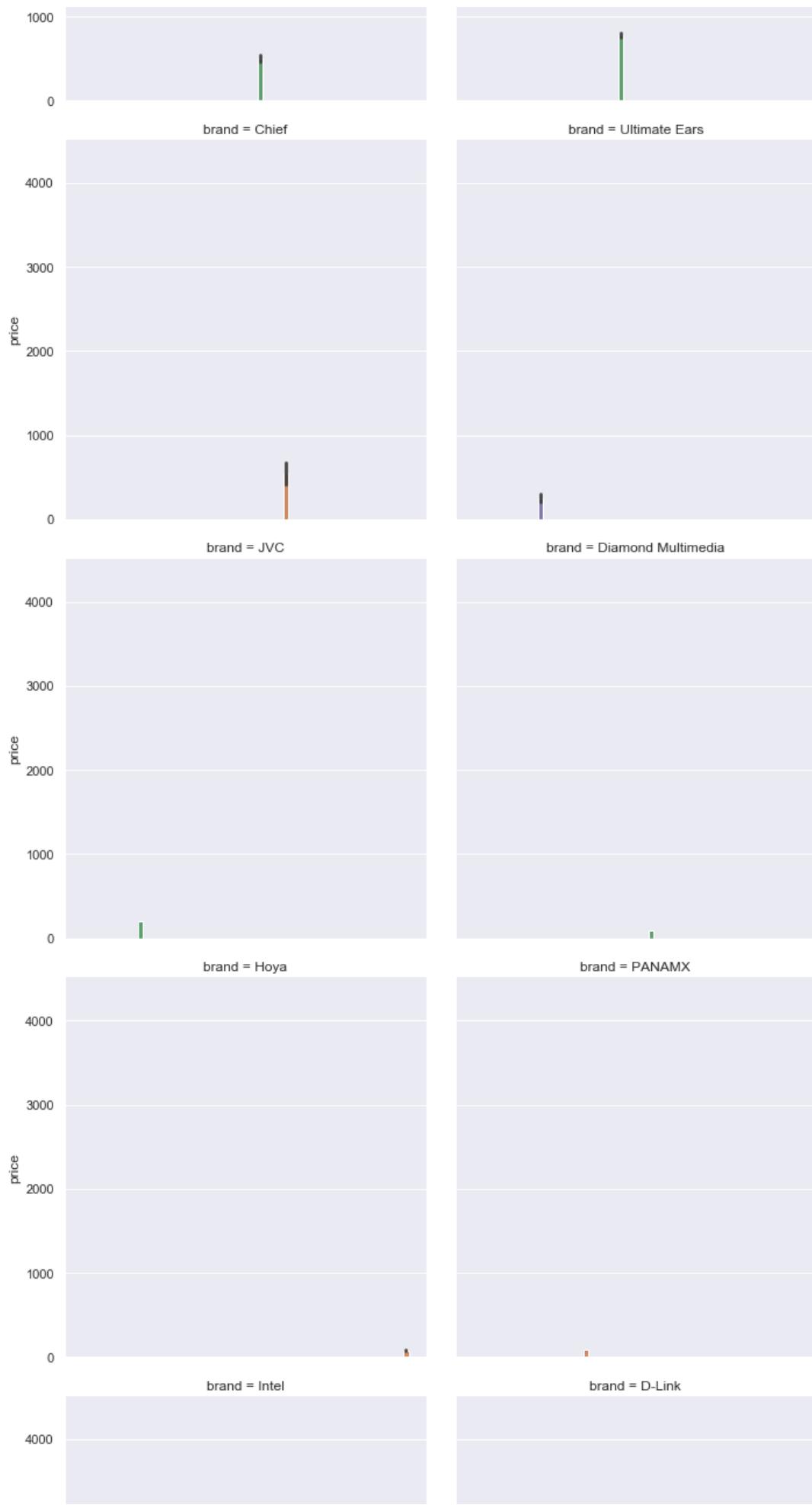


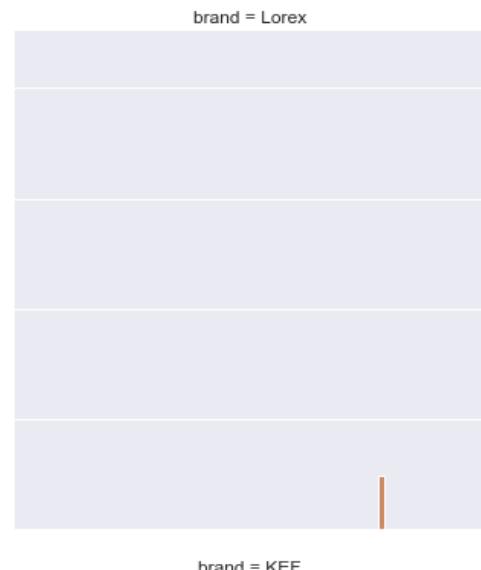
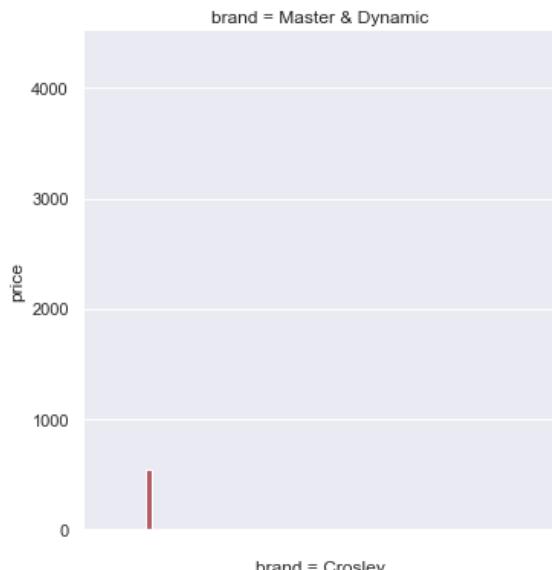
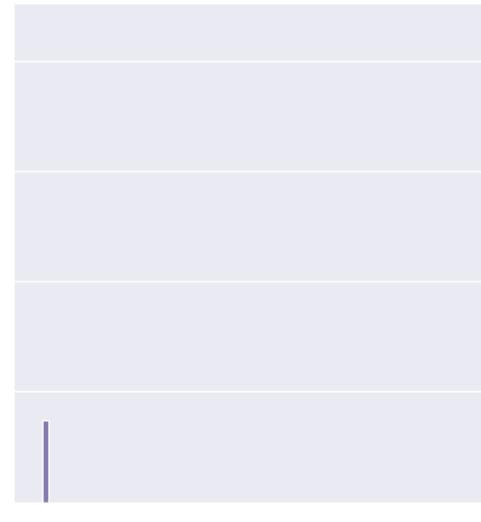
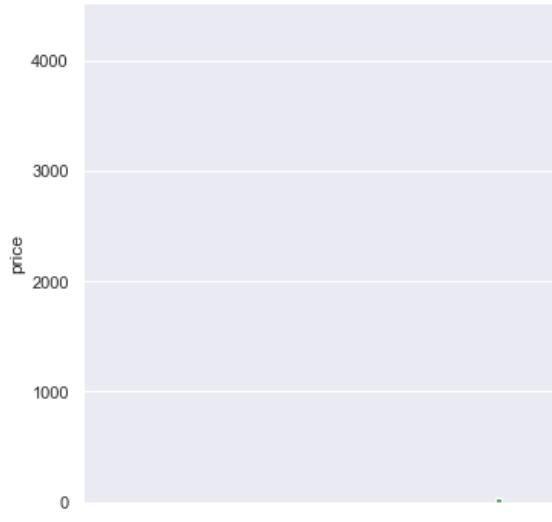
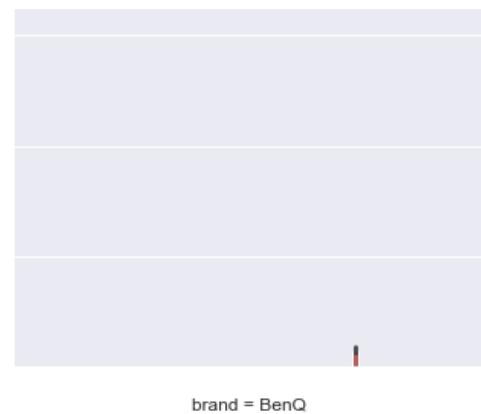
brand = FUGOO

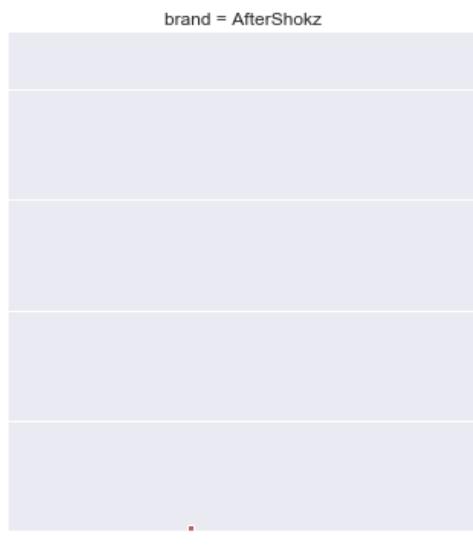
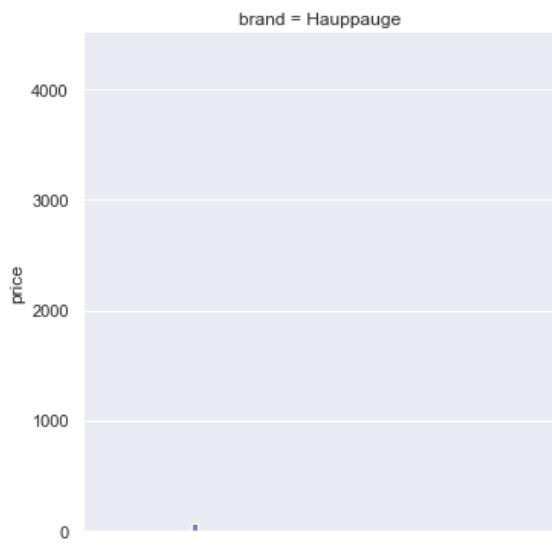
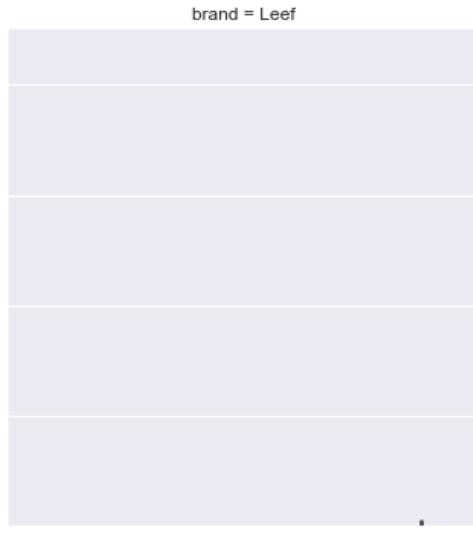
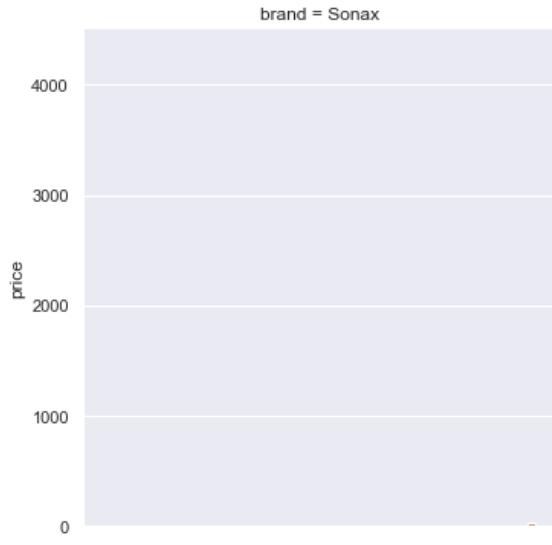
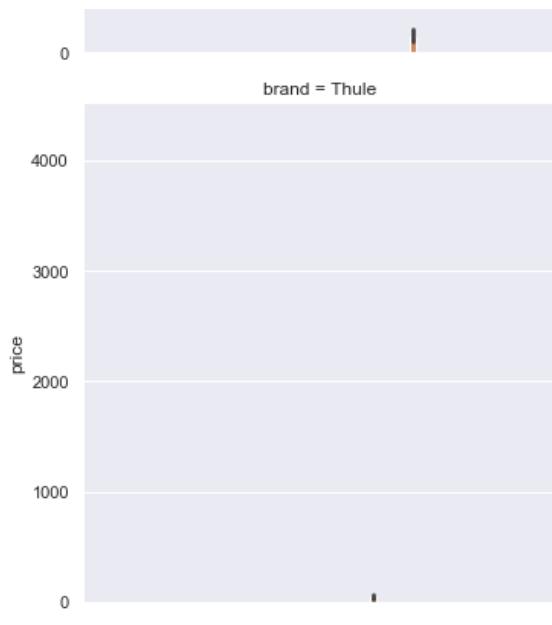


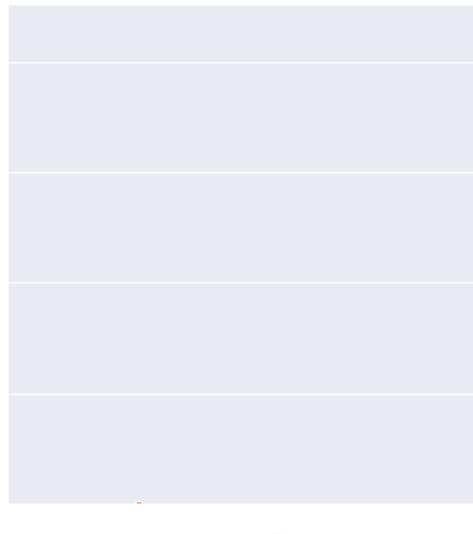
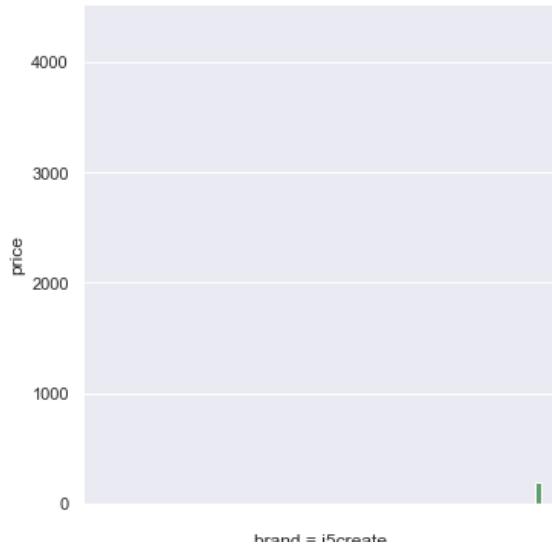
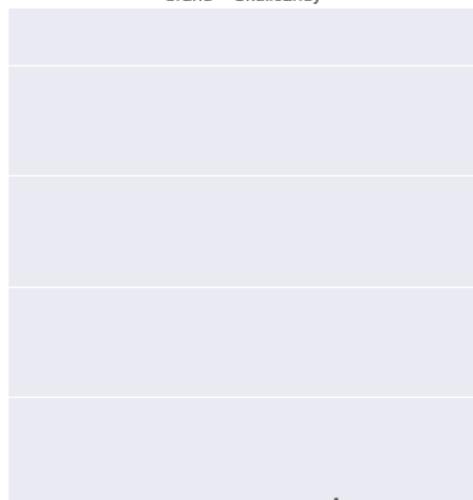
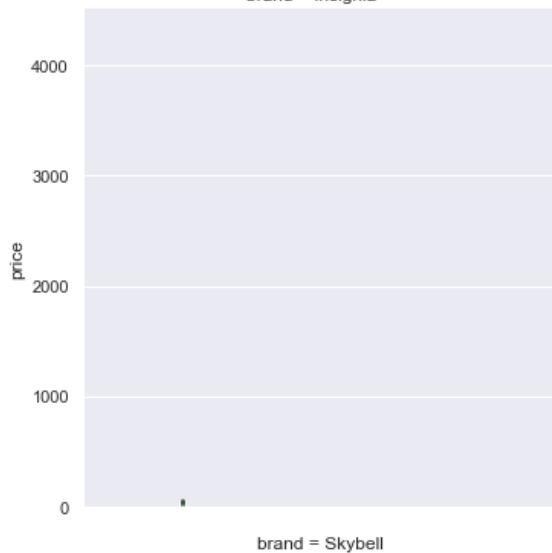
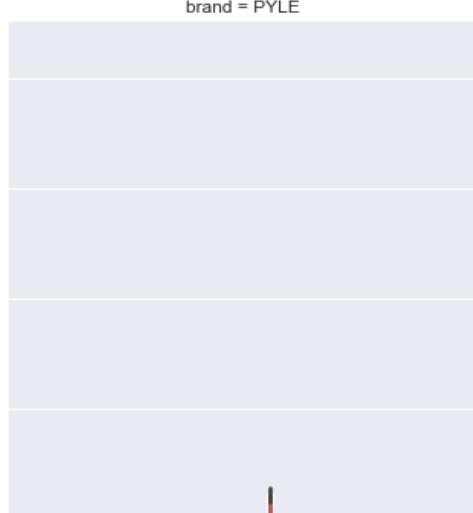
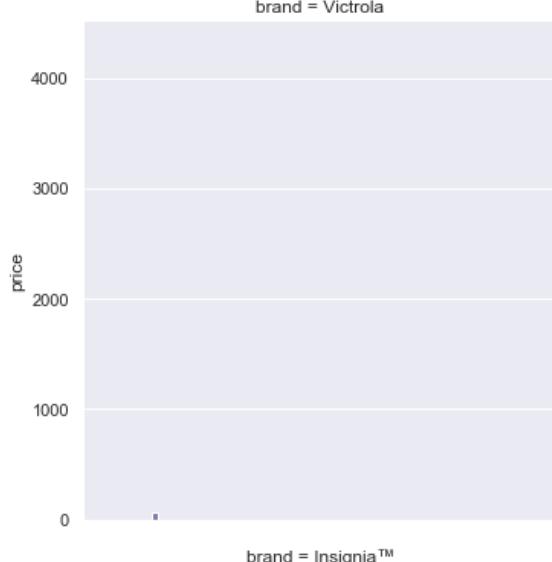
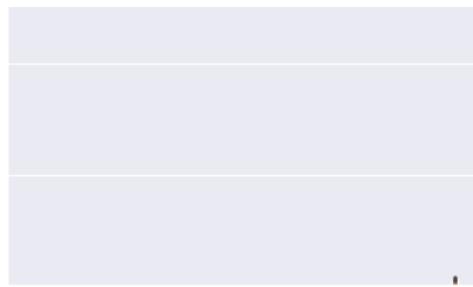


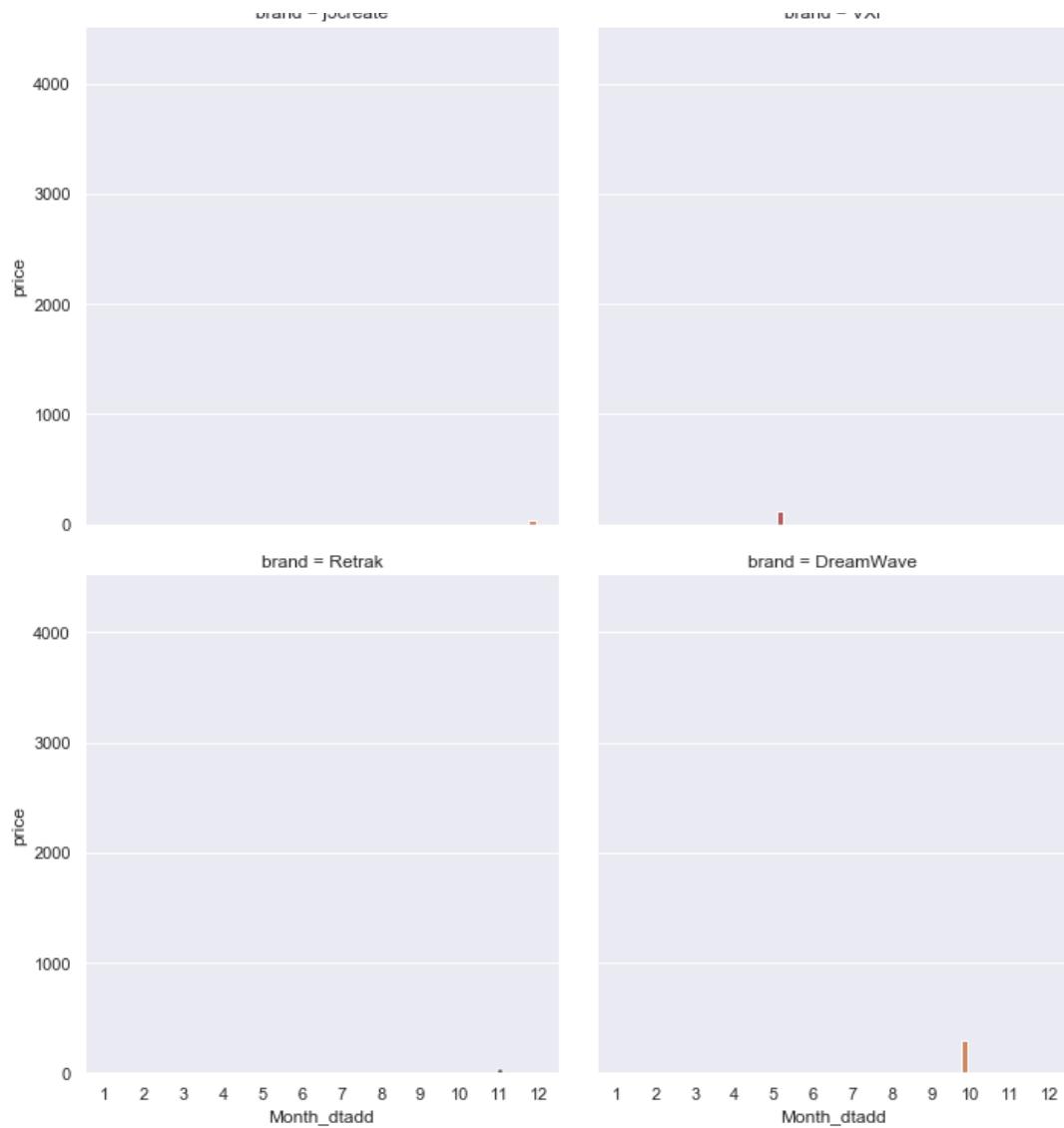










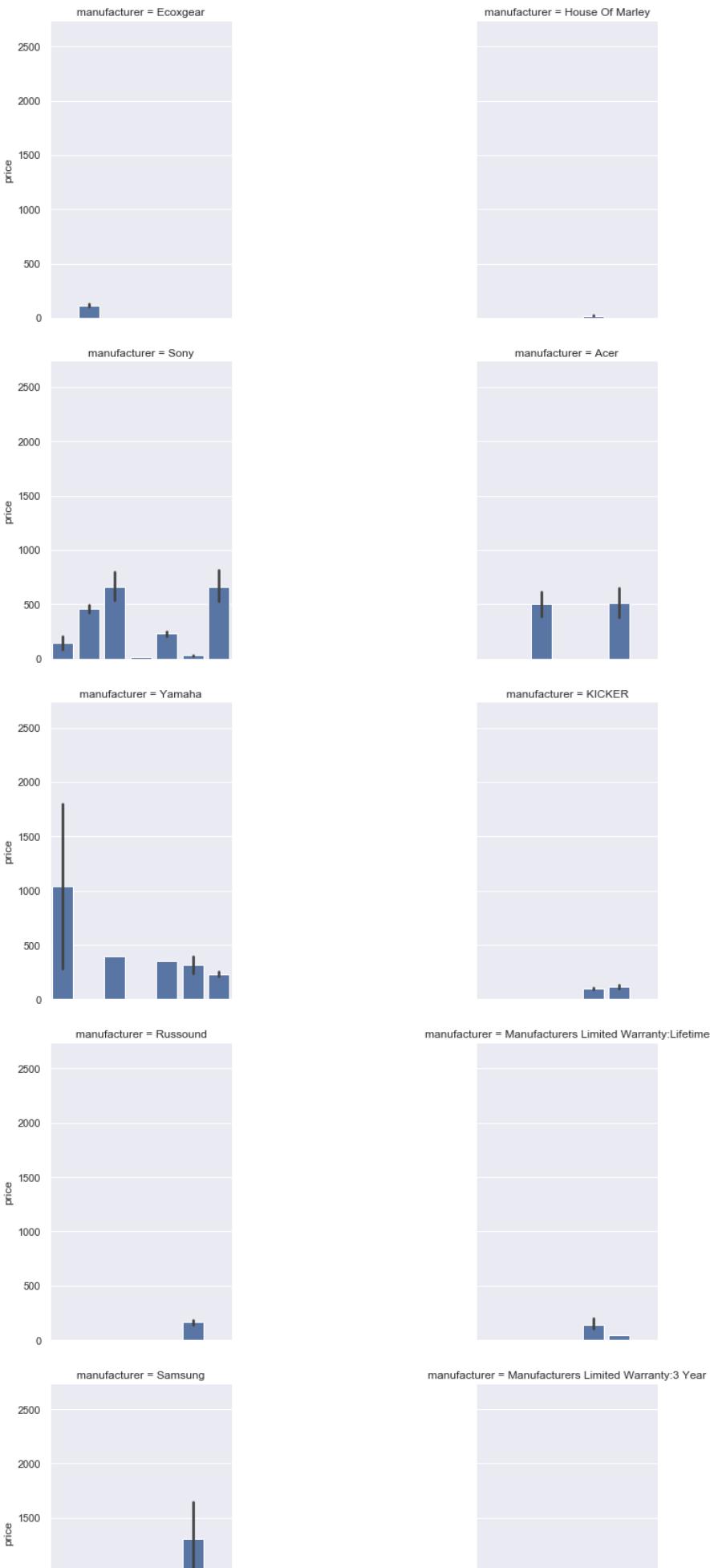


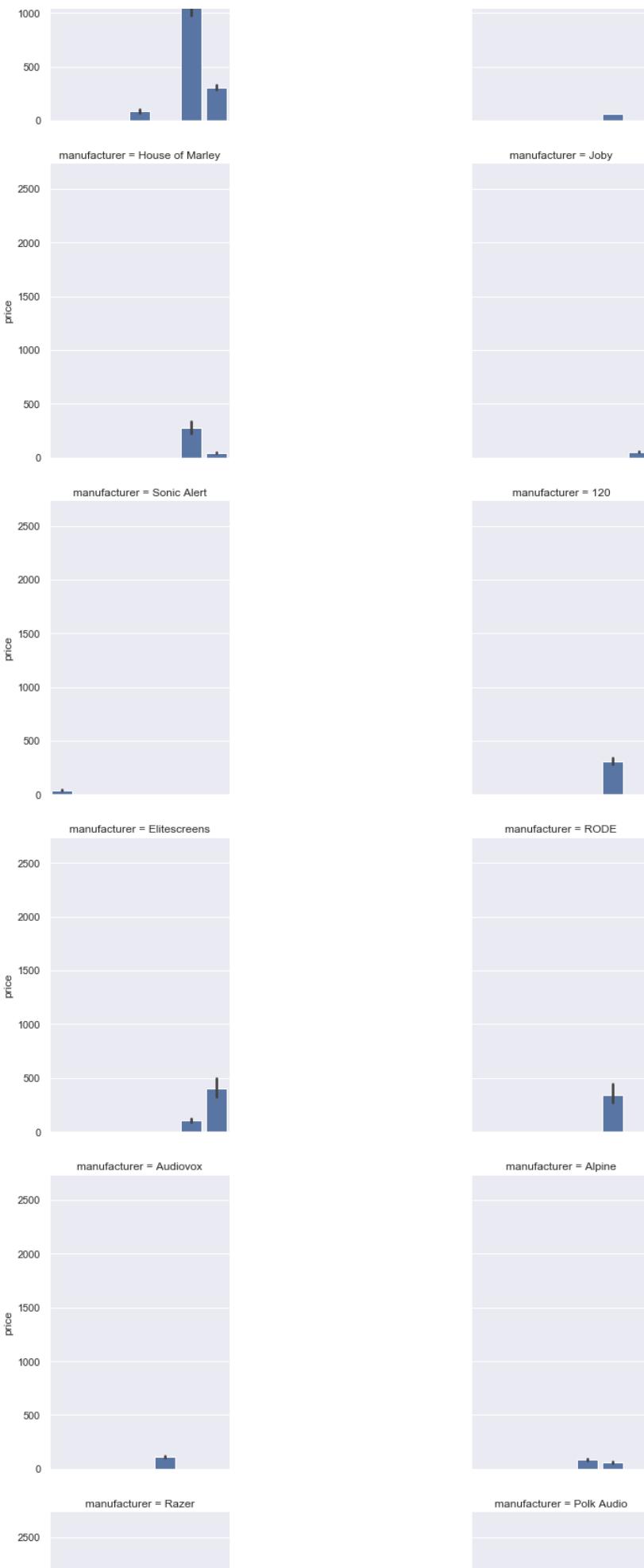
## Observation

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

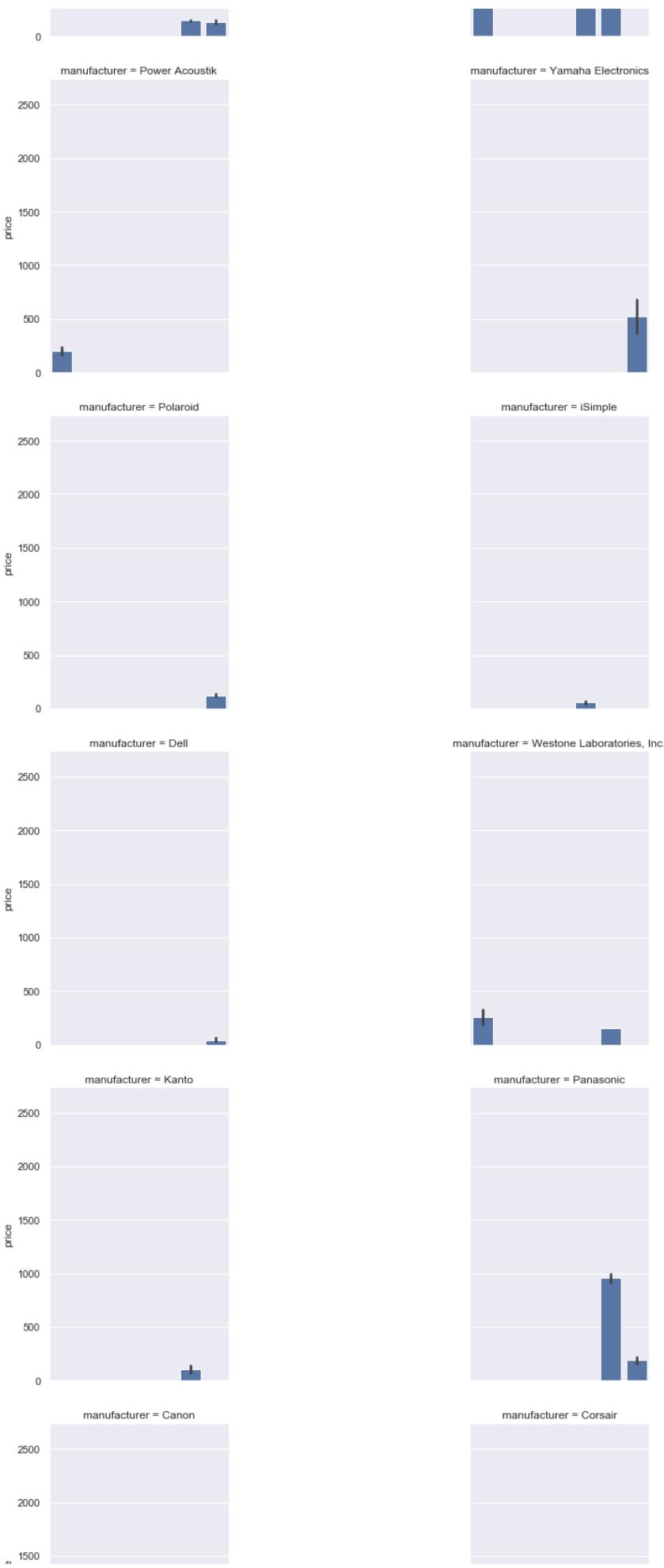
I will plot manufacturer with dateUpdated

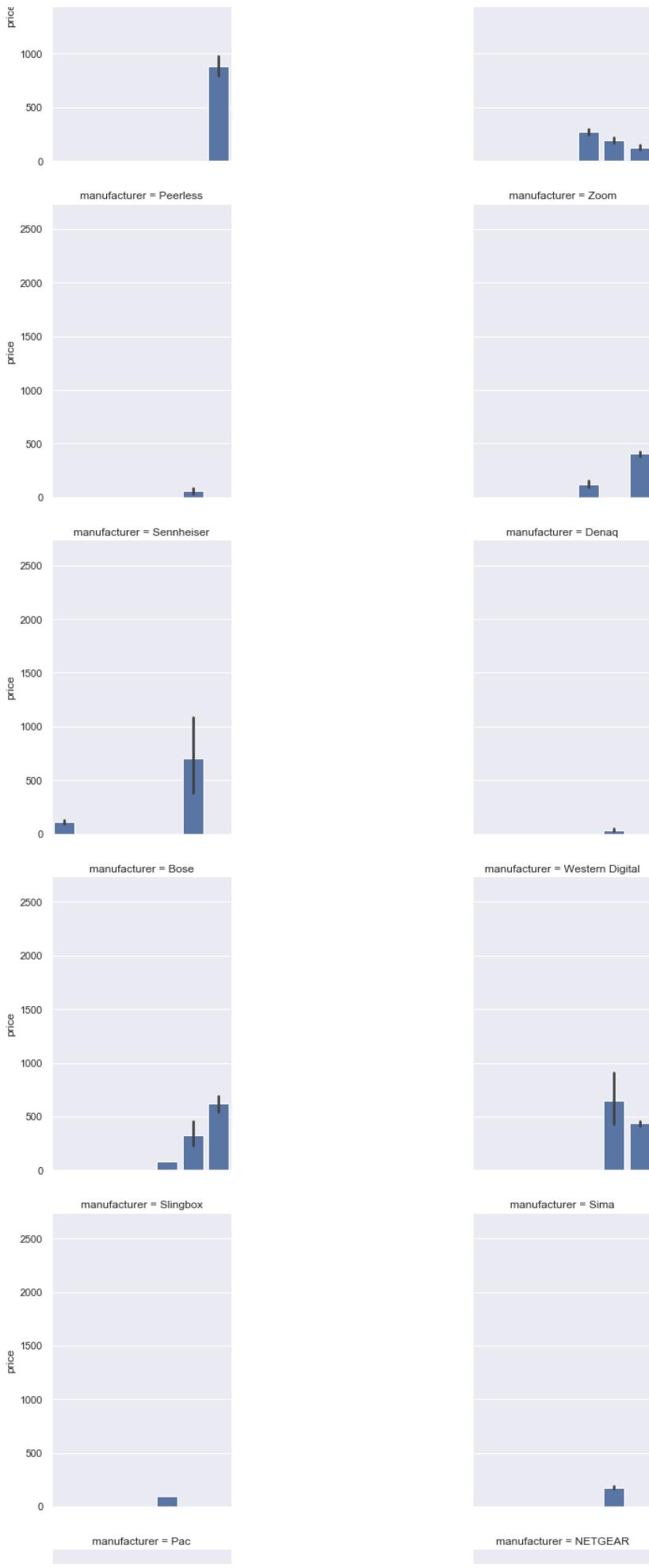
```
In [75]: sns.set(style='darkgrid', font_scale=1.0)
sns.catplot(x='Month_updated',
             y='price',hue='Year_updated',
             col='manufacturer',
             col_wrap=2,
             data=dt_electronics,
             kind='bar')
plt.savefig('Dateupdated_manufacturer.png')
```

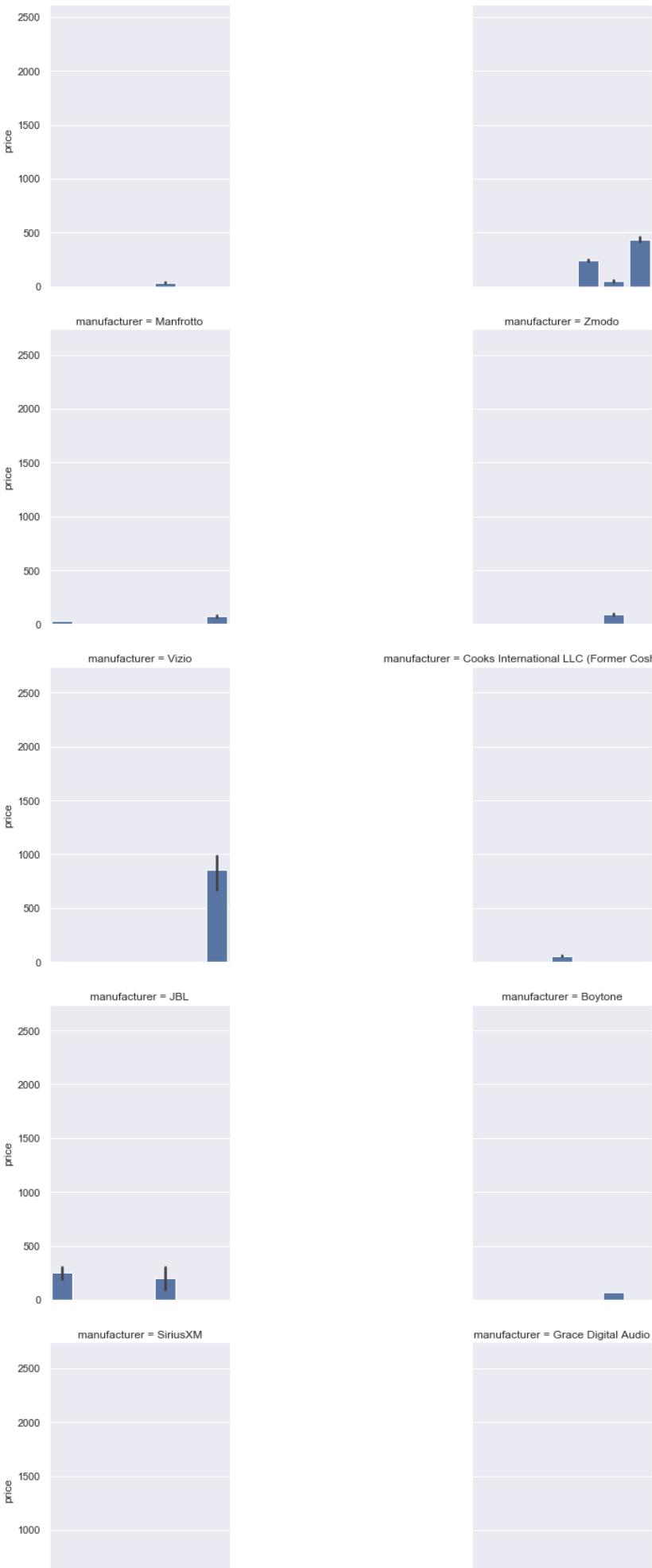


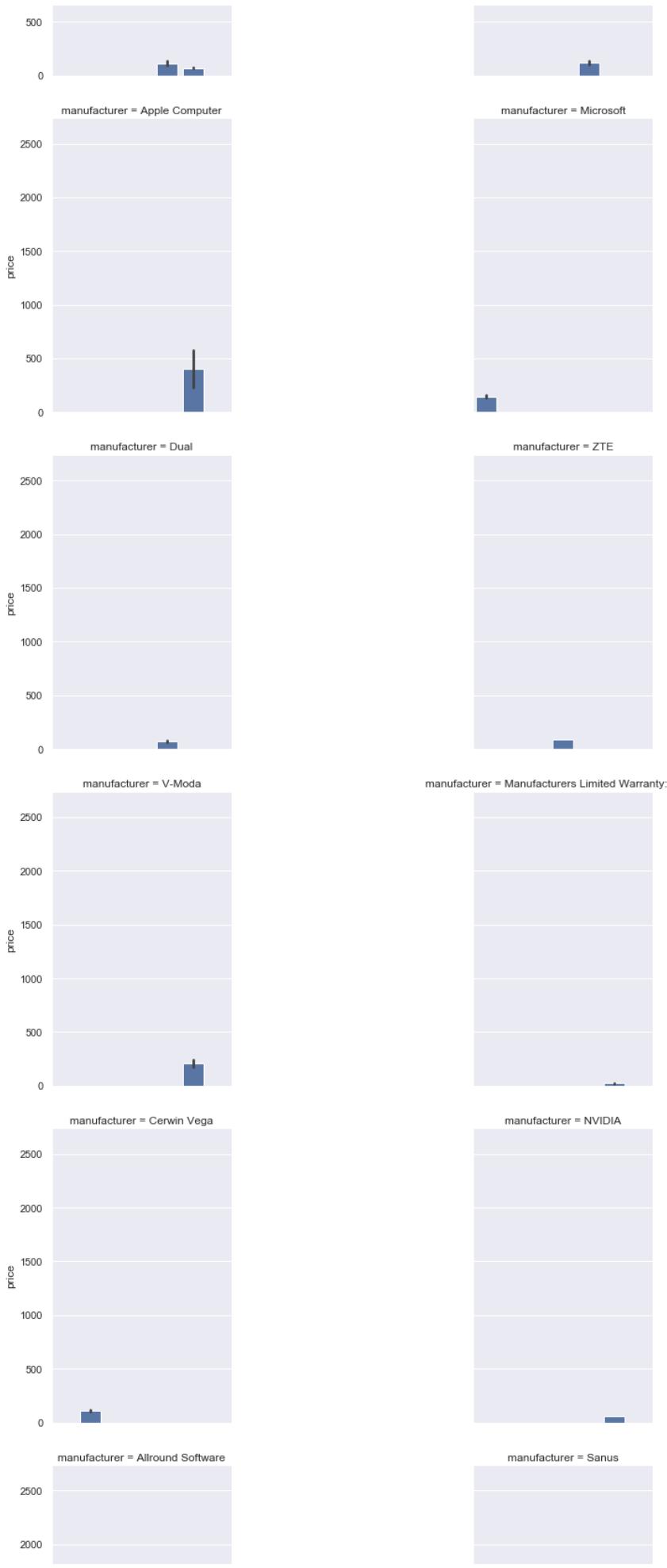


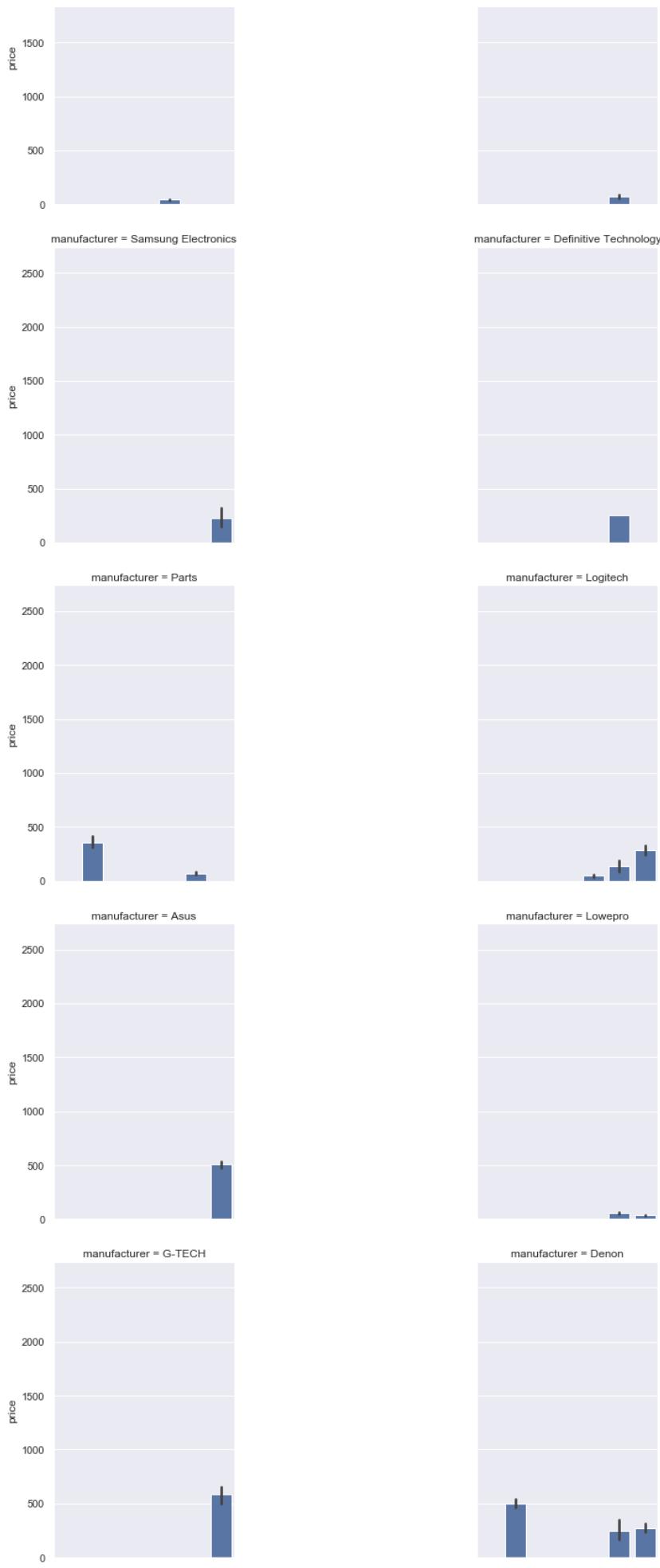


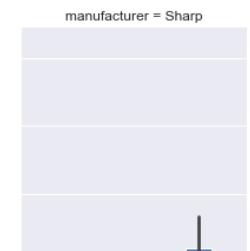
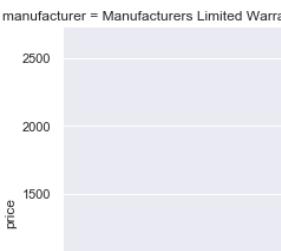
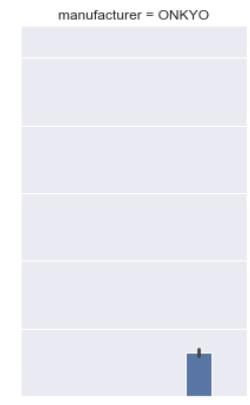
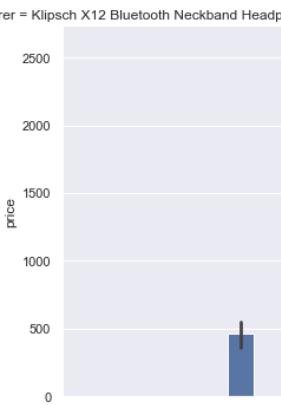
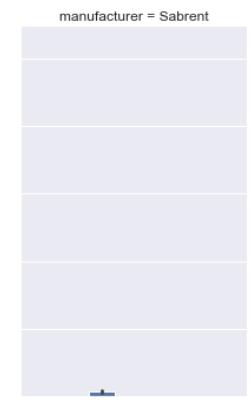
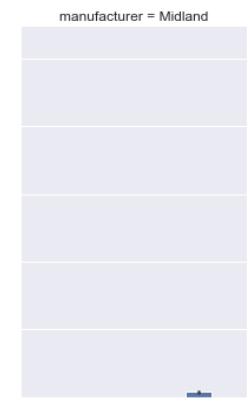
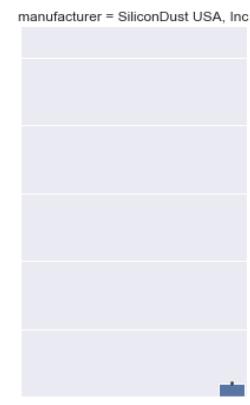


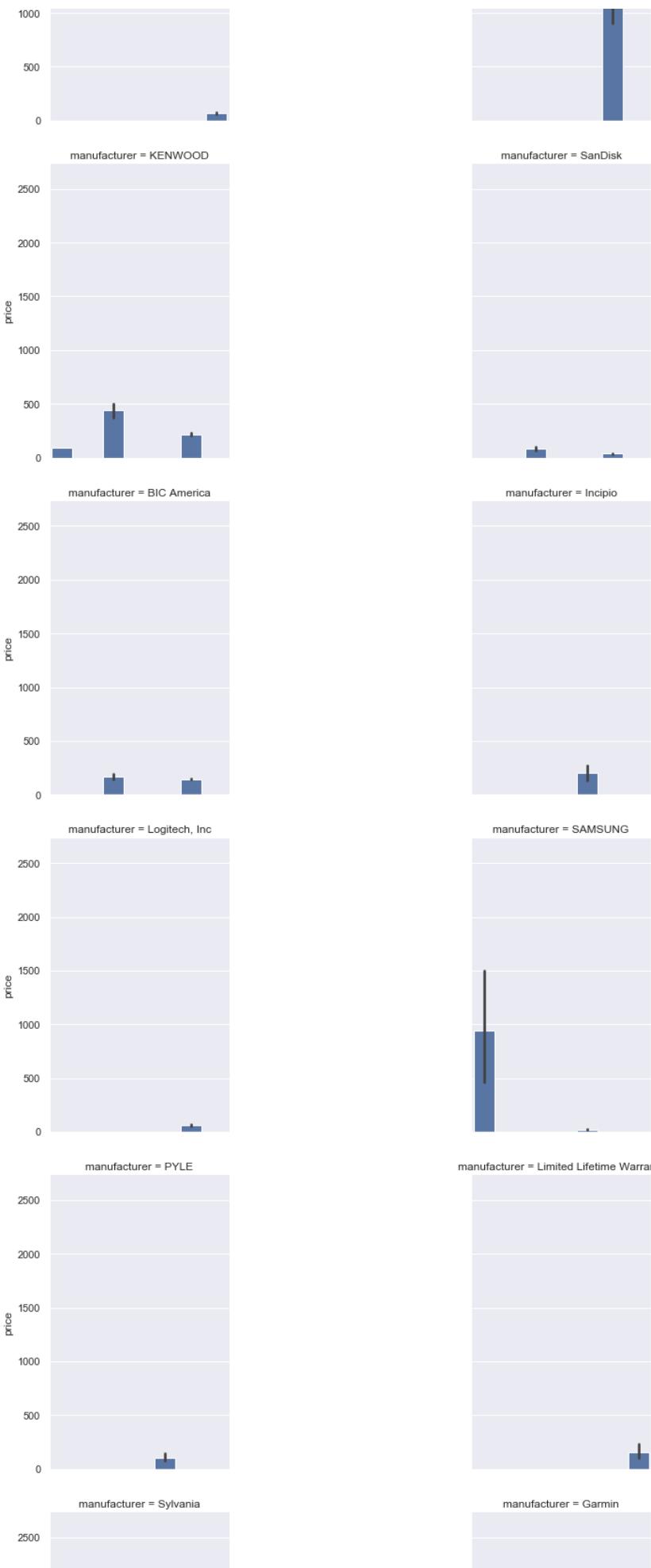


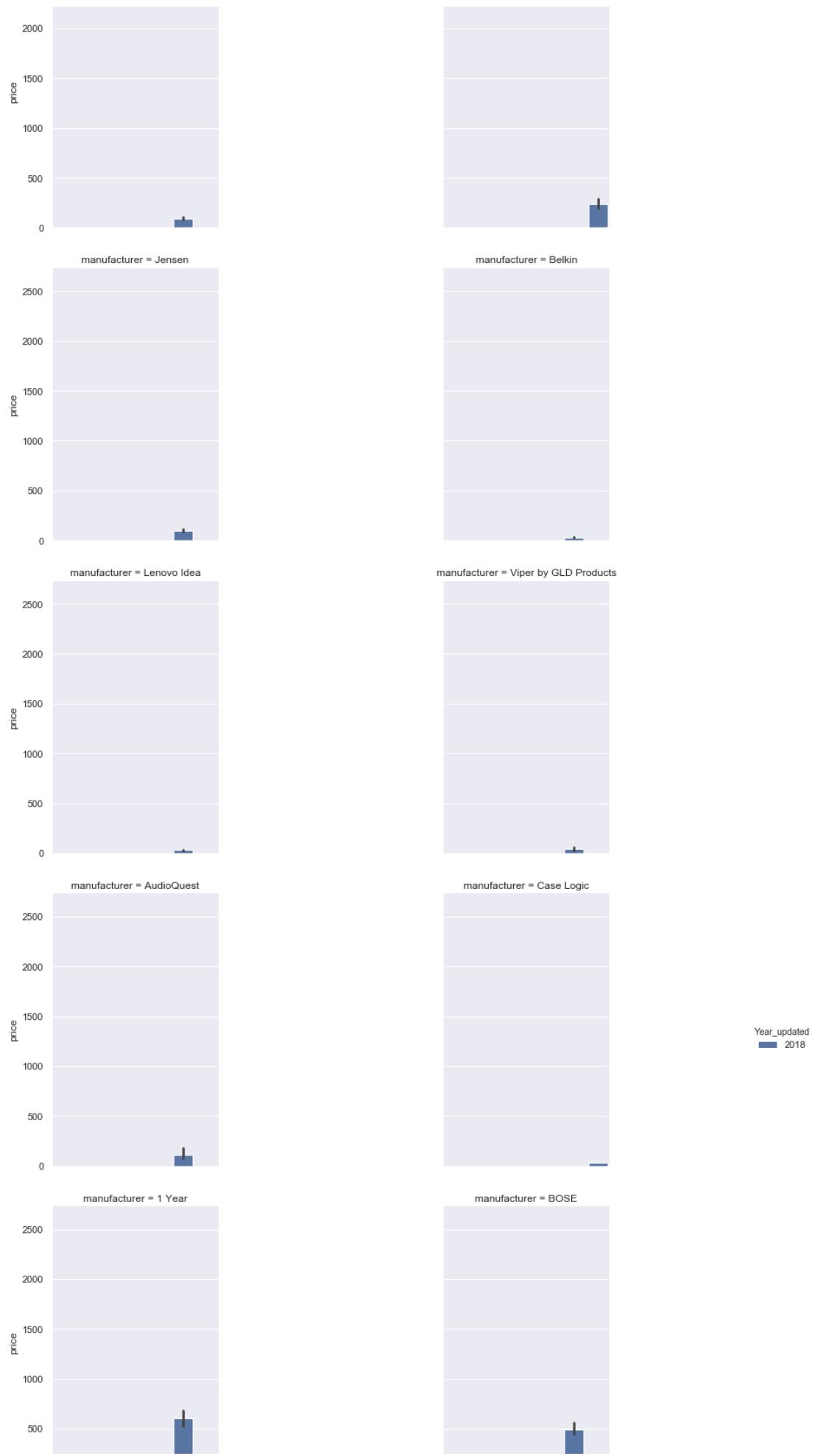


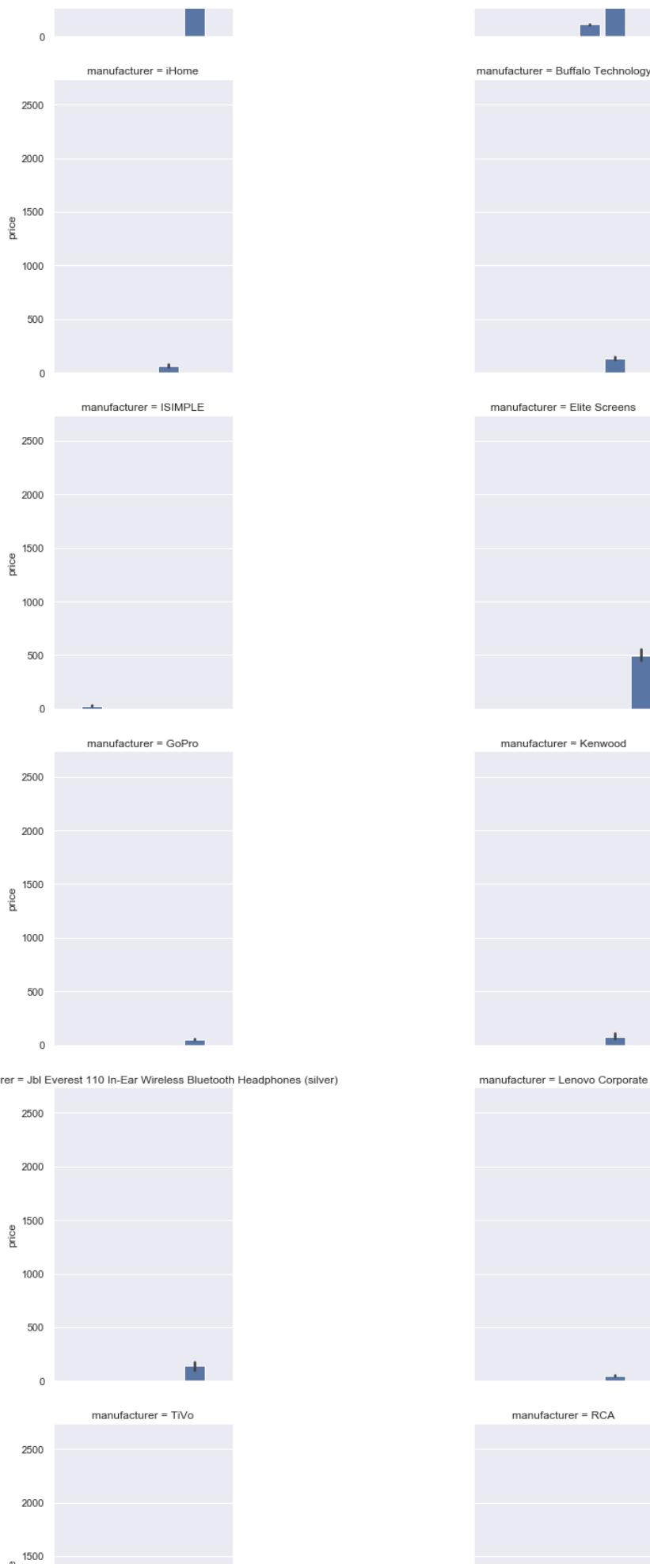




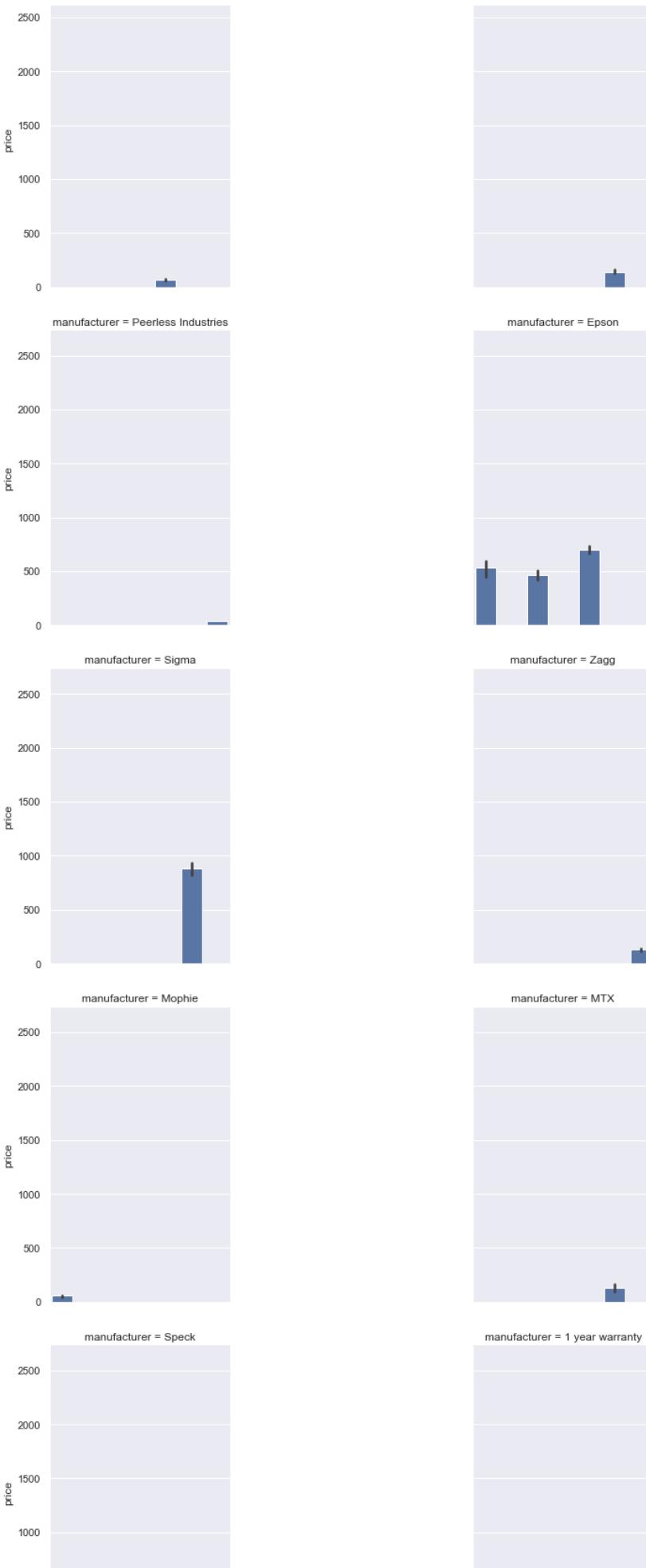




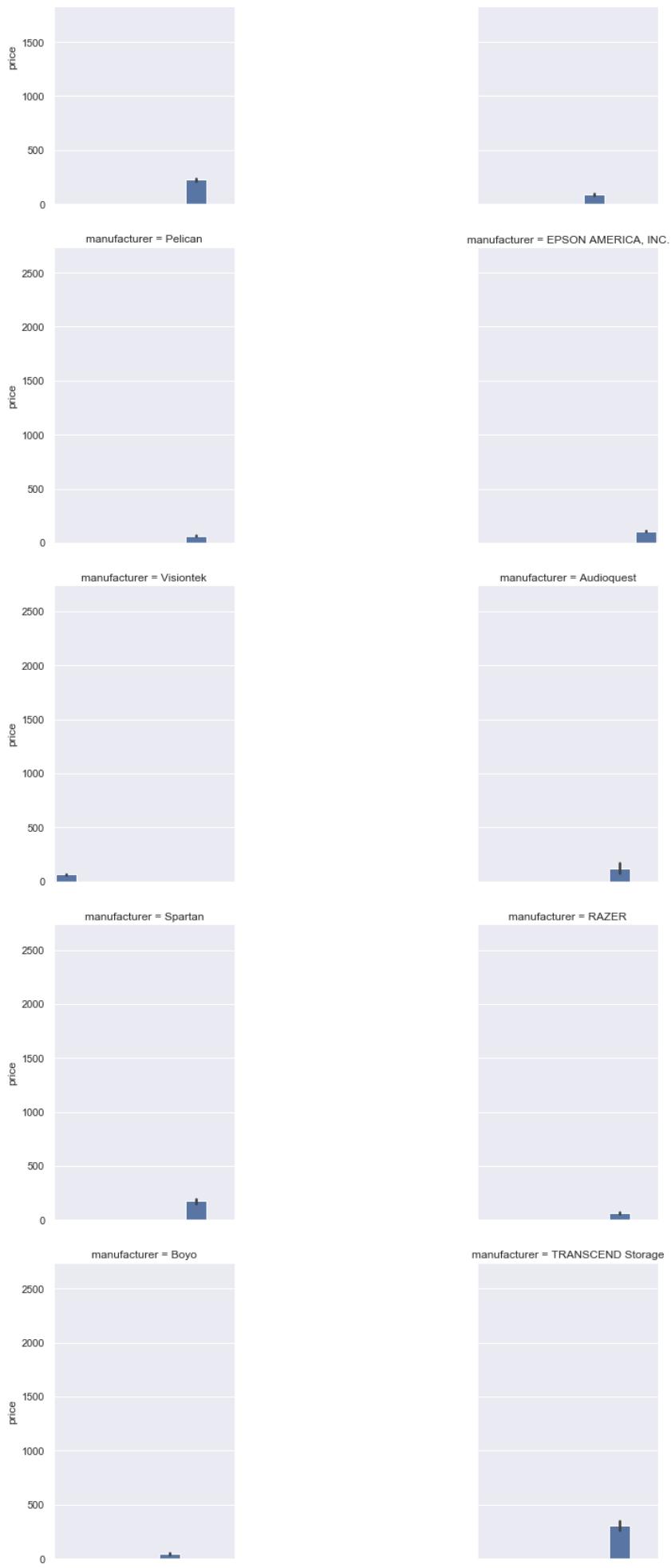


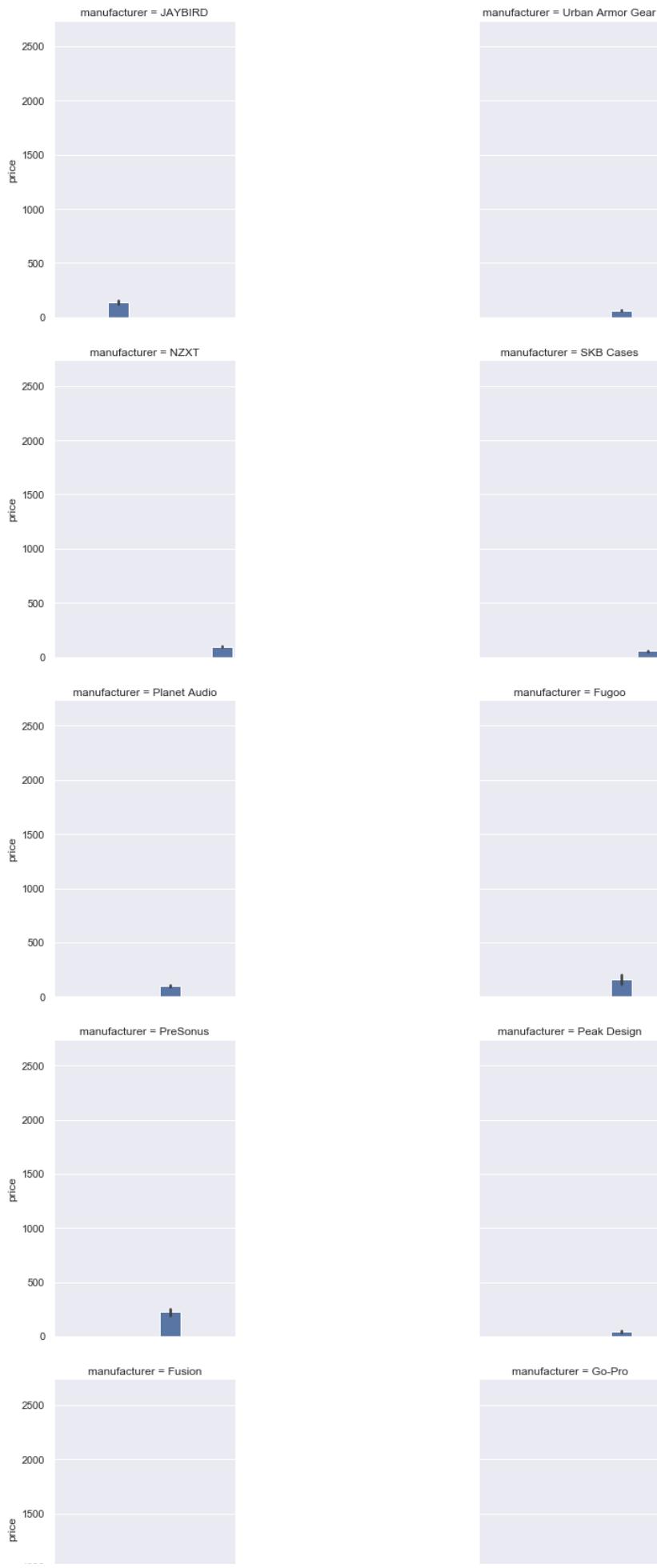


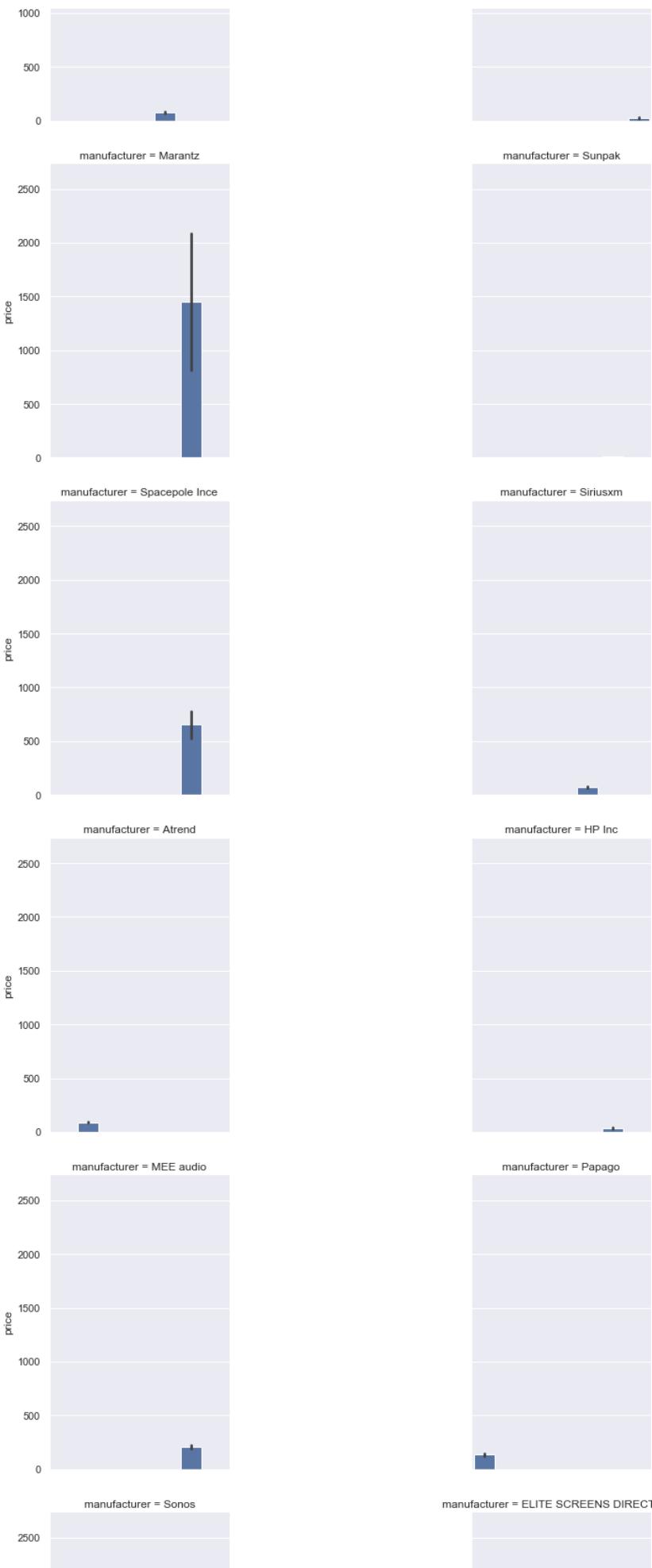


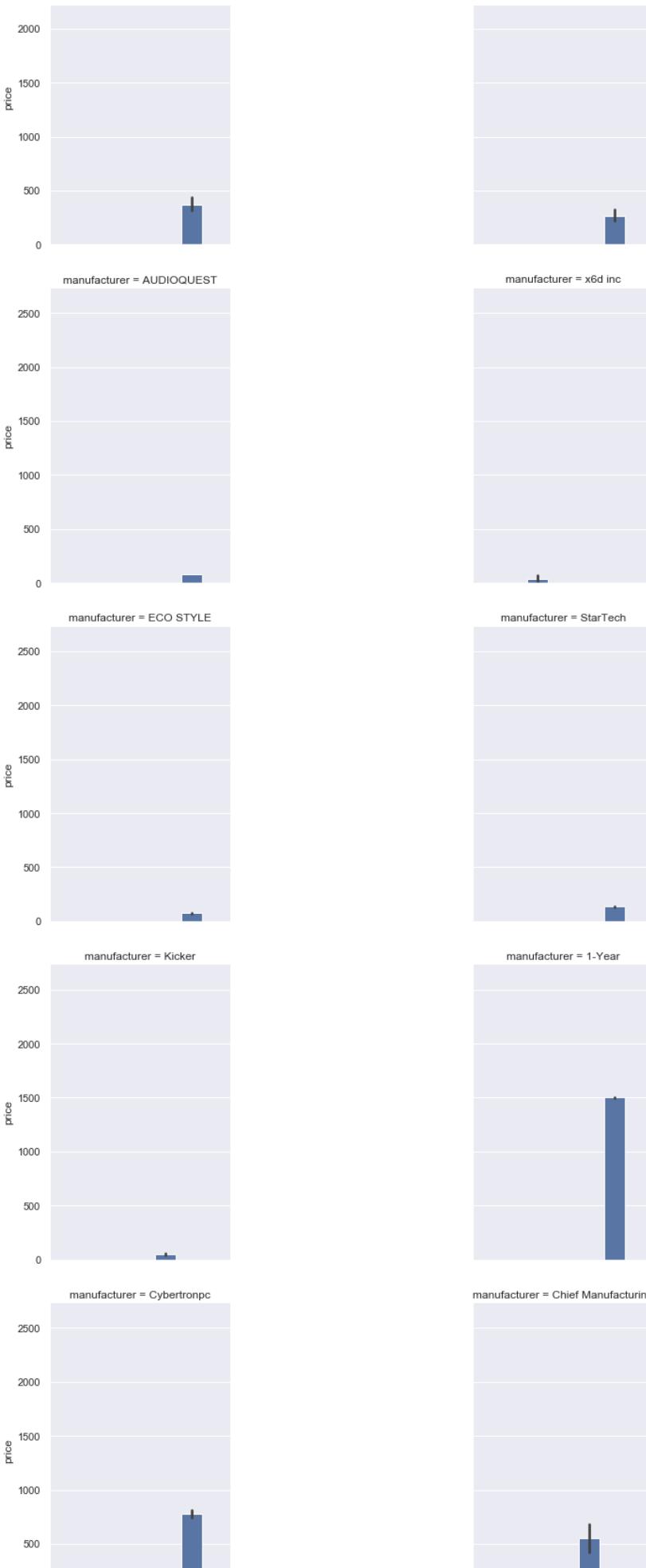


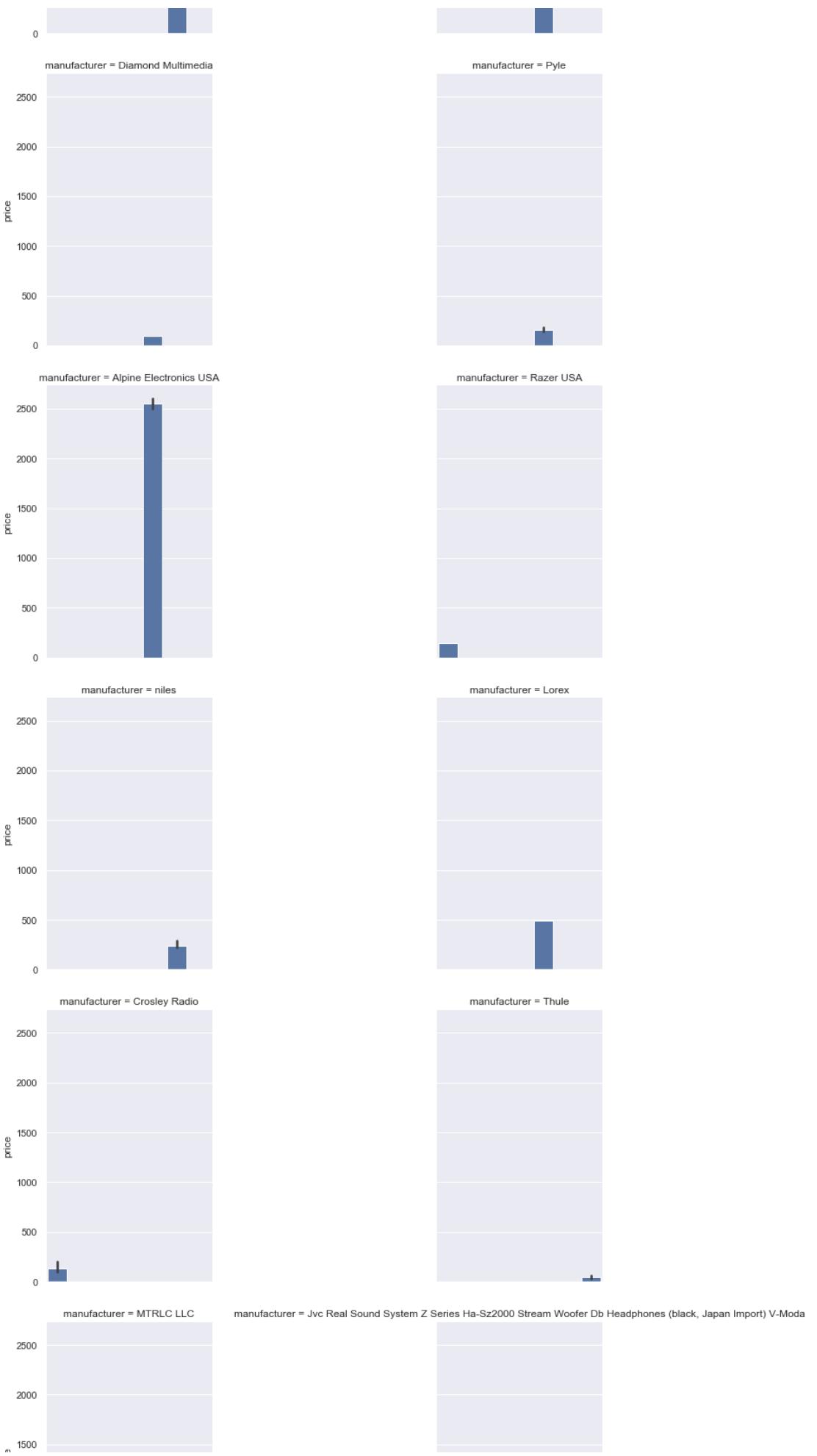




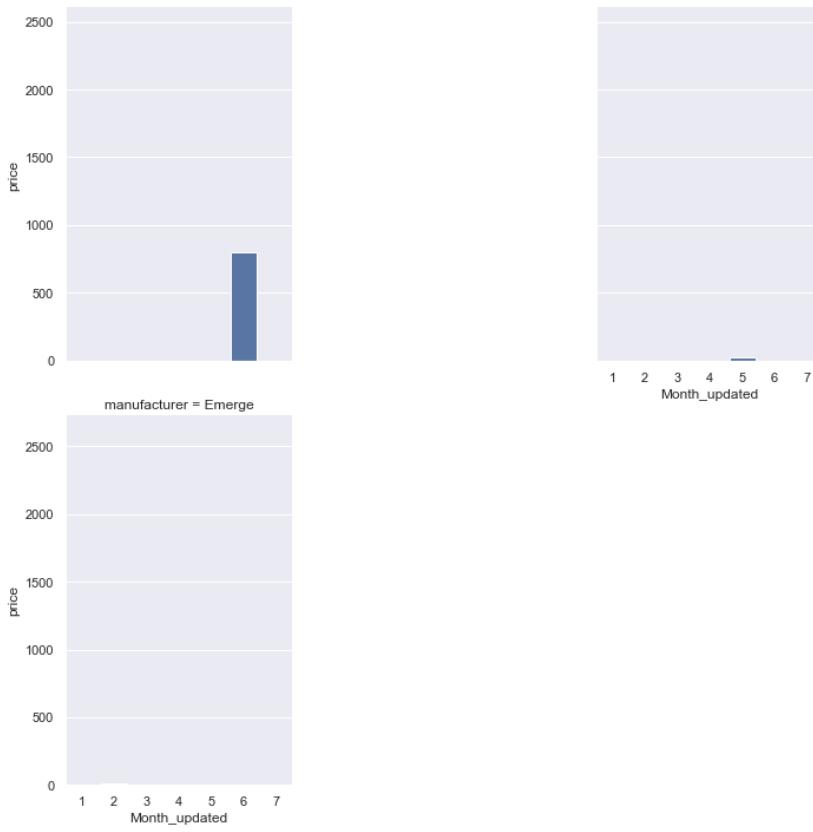










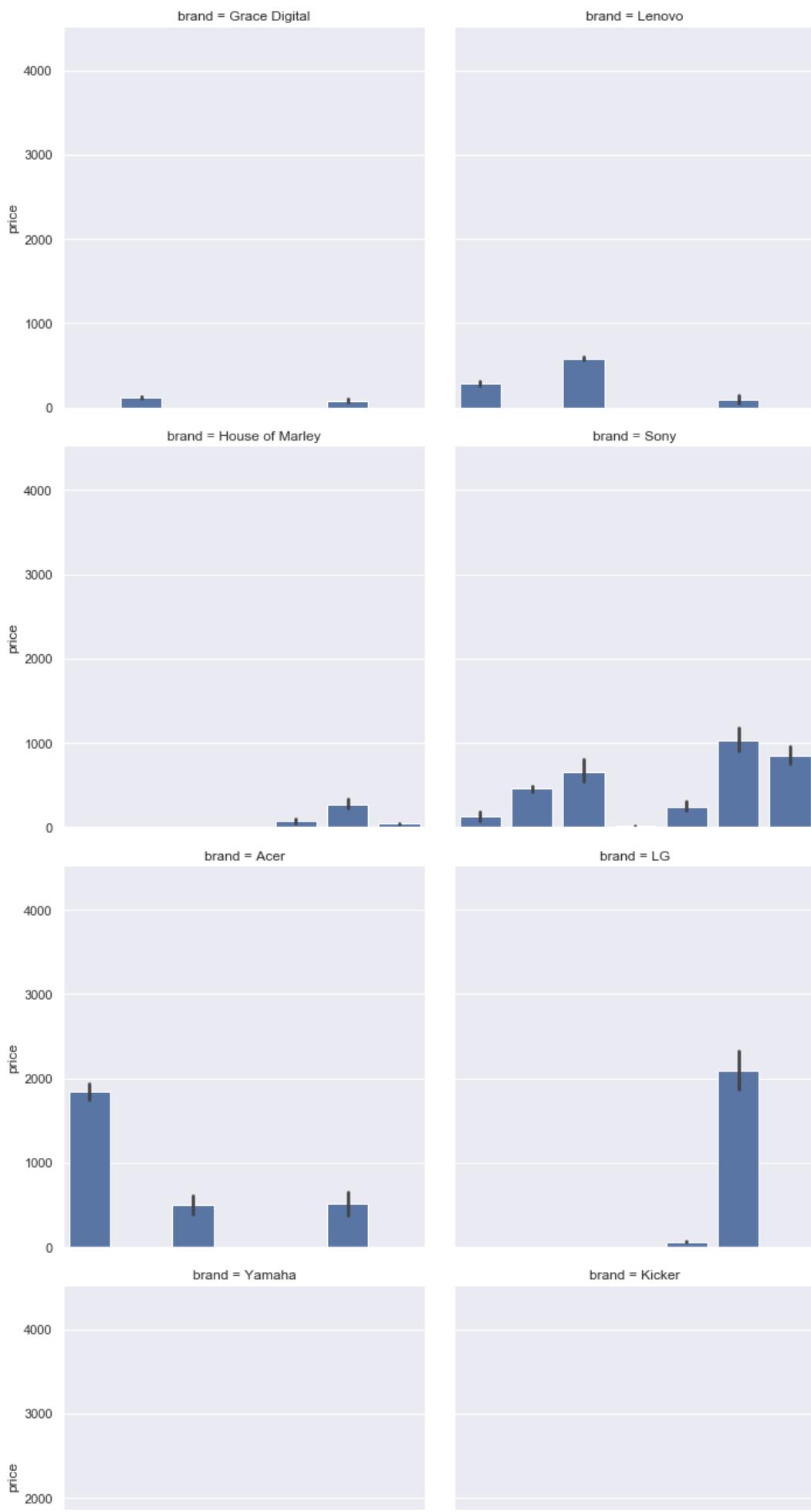


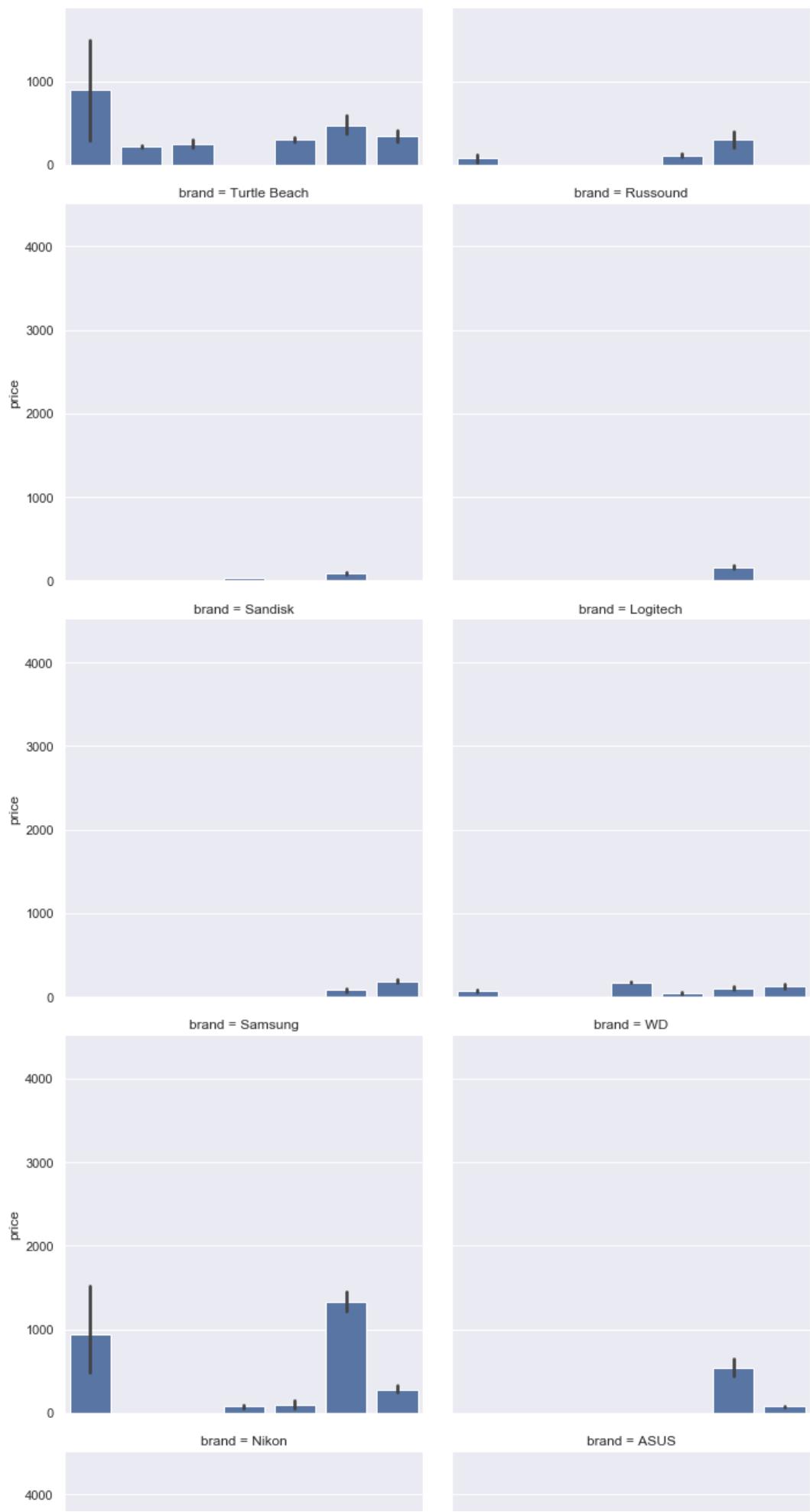
## Observation

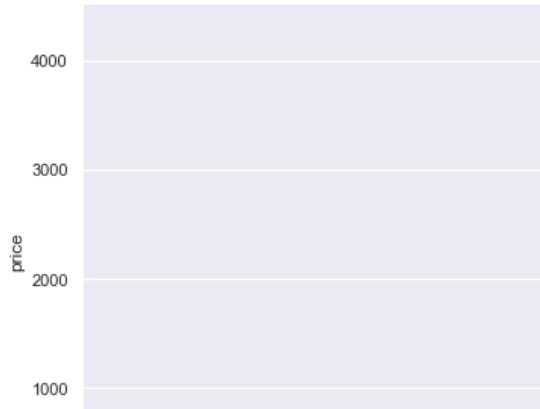
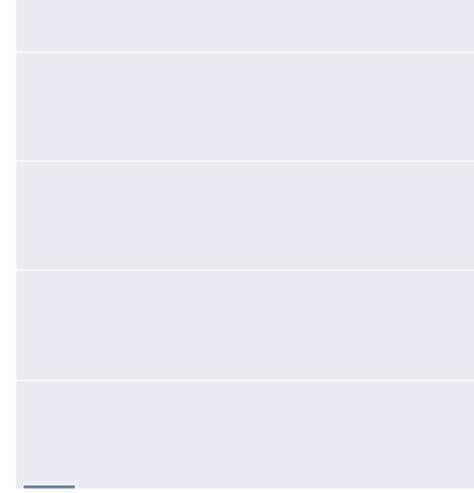
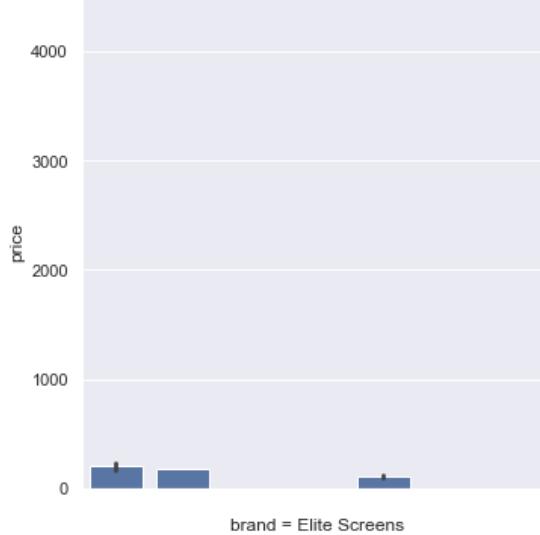
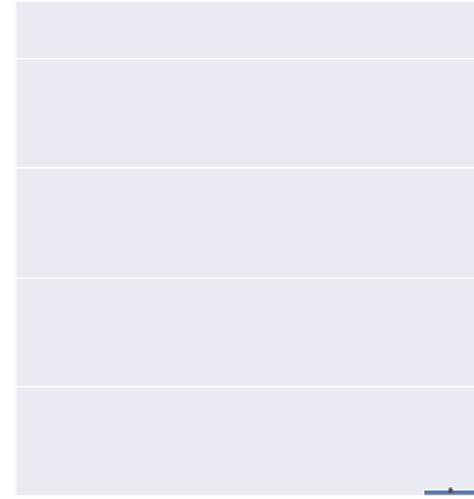
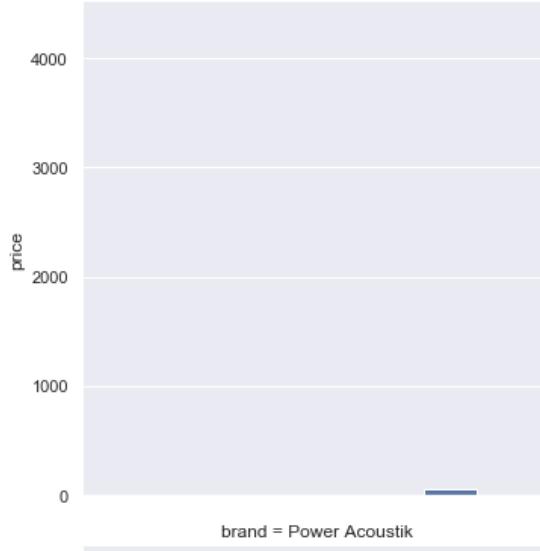
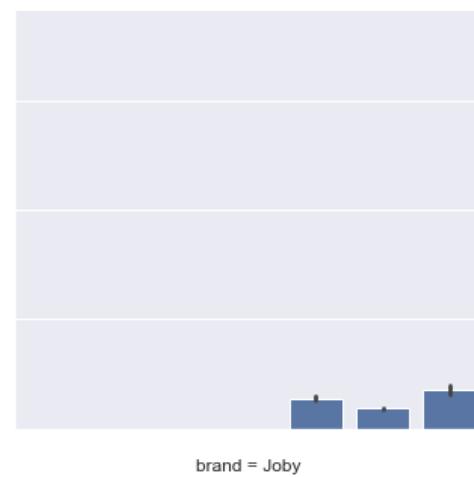
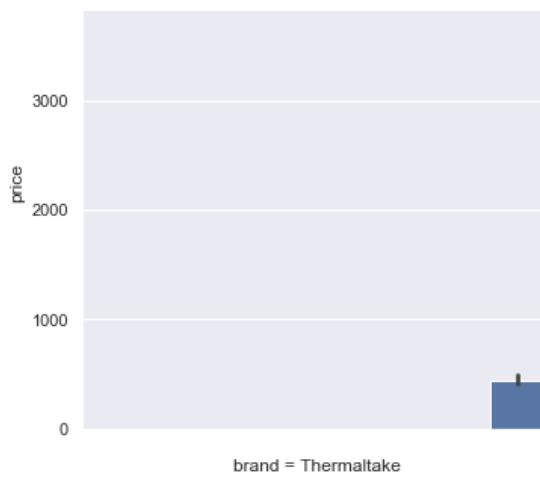
From this graph I can conclude Apple has more price than other category  
I can also conclude that in 1st and 5th month of 2018 price is highest

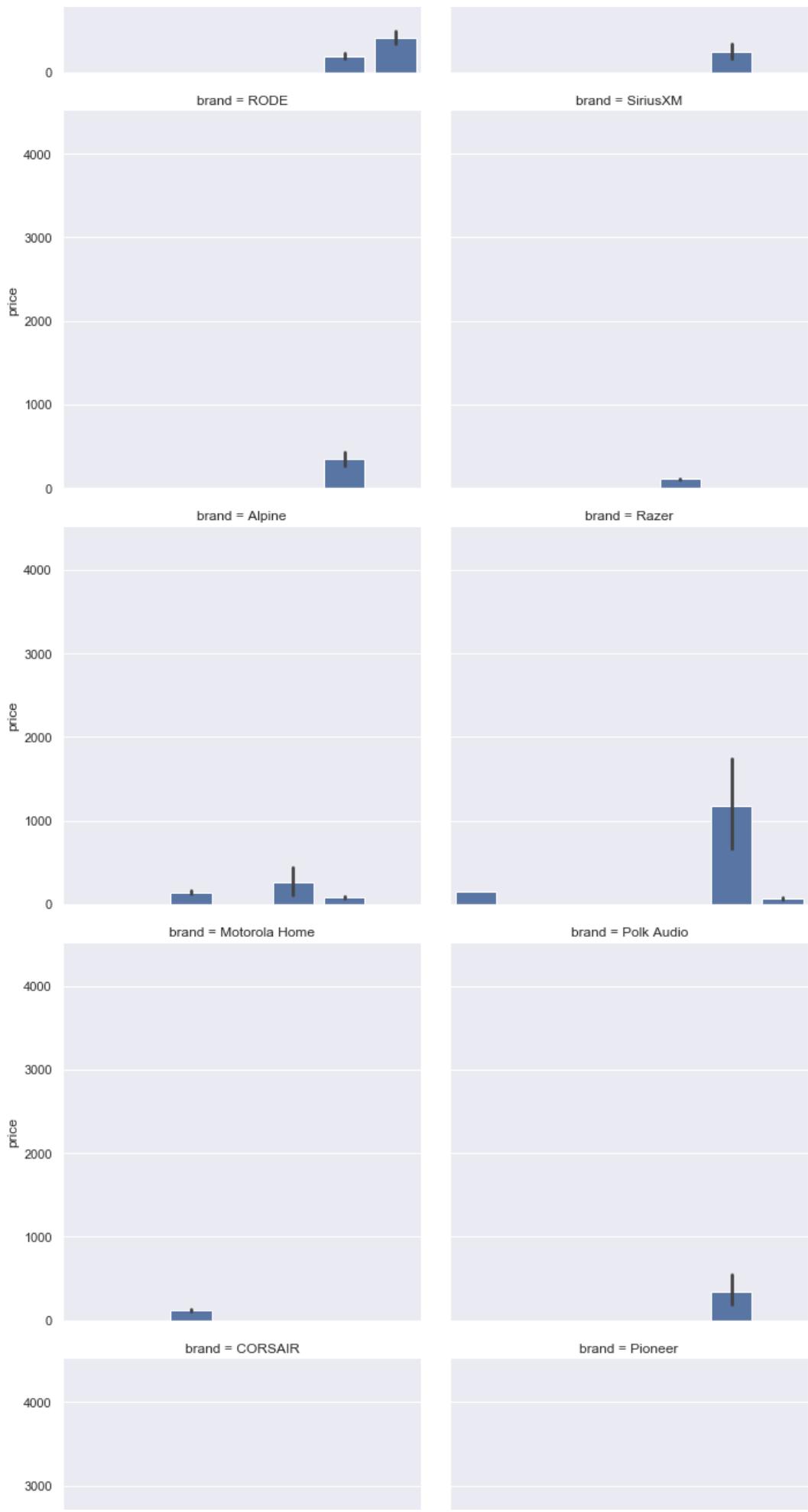
I will plot band with dateupdated

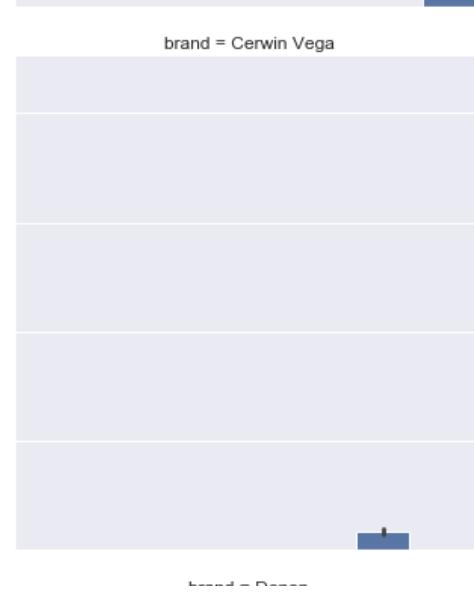
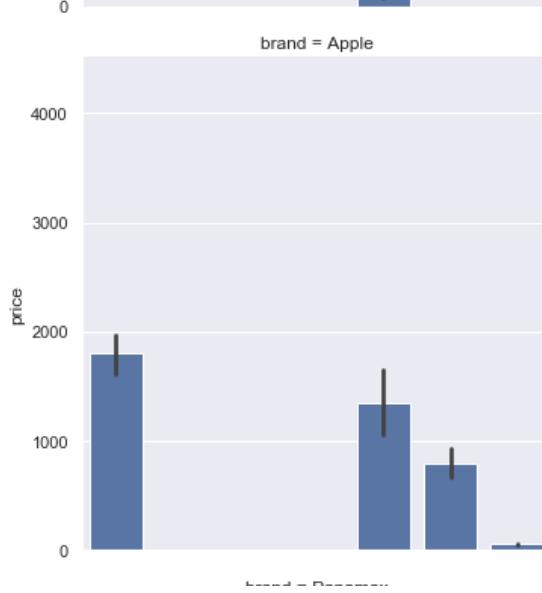
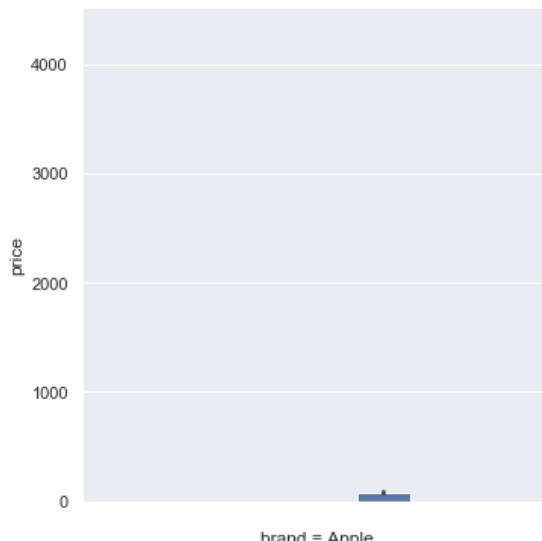
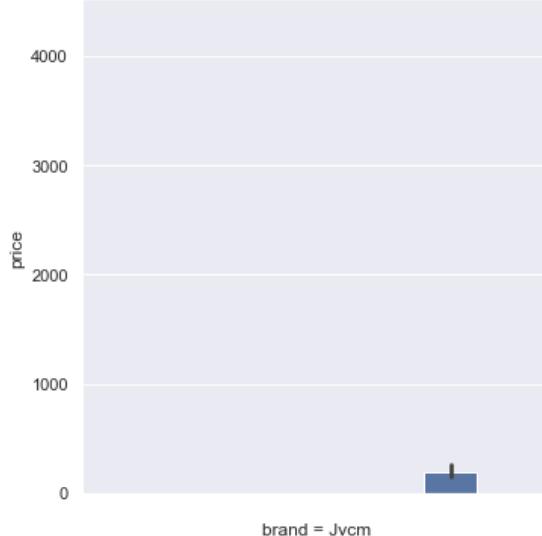
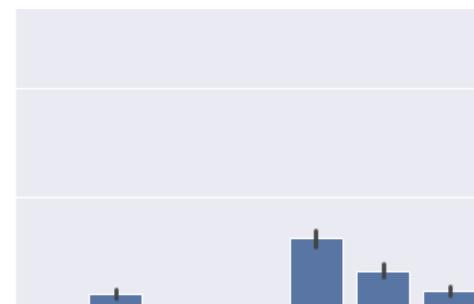
```
In [67]: sns.set(style='darkgrid', font_scale=1.0)
sns.catplot(x='Month_updated',
             y='price',hue='Year_updated',
             col='brand',
             col_wrap=2,
             data=dt_electronics,
             kind='bar')
plt.savefig('Dateupdated_brand.png')
```

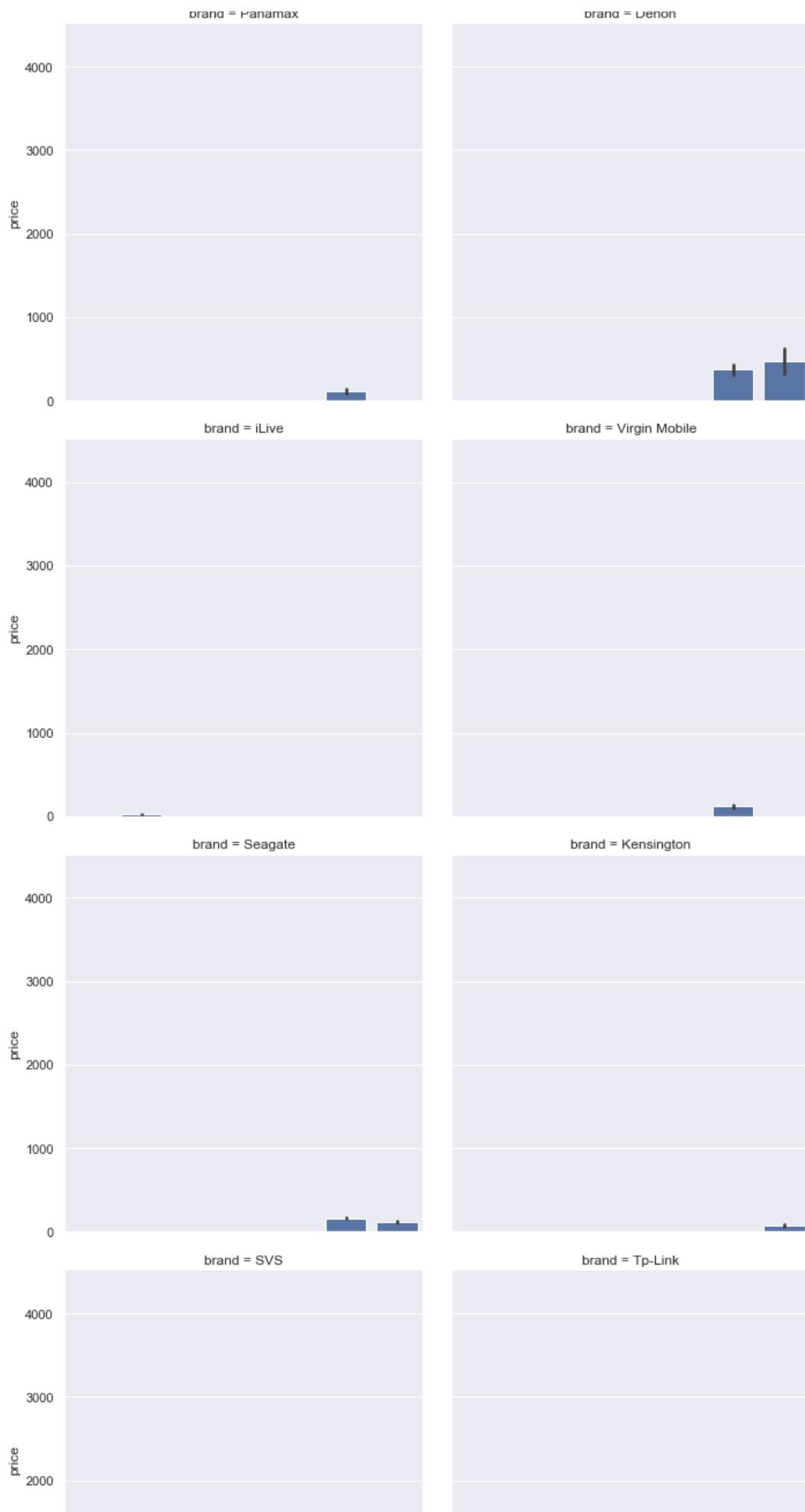


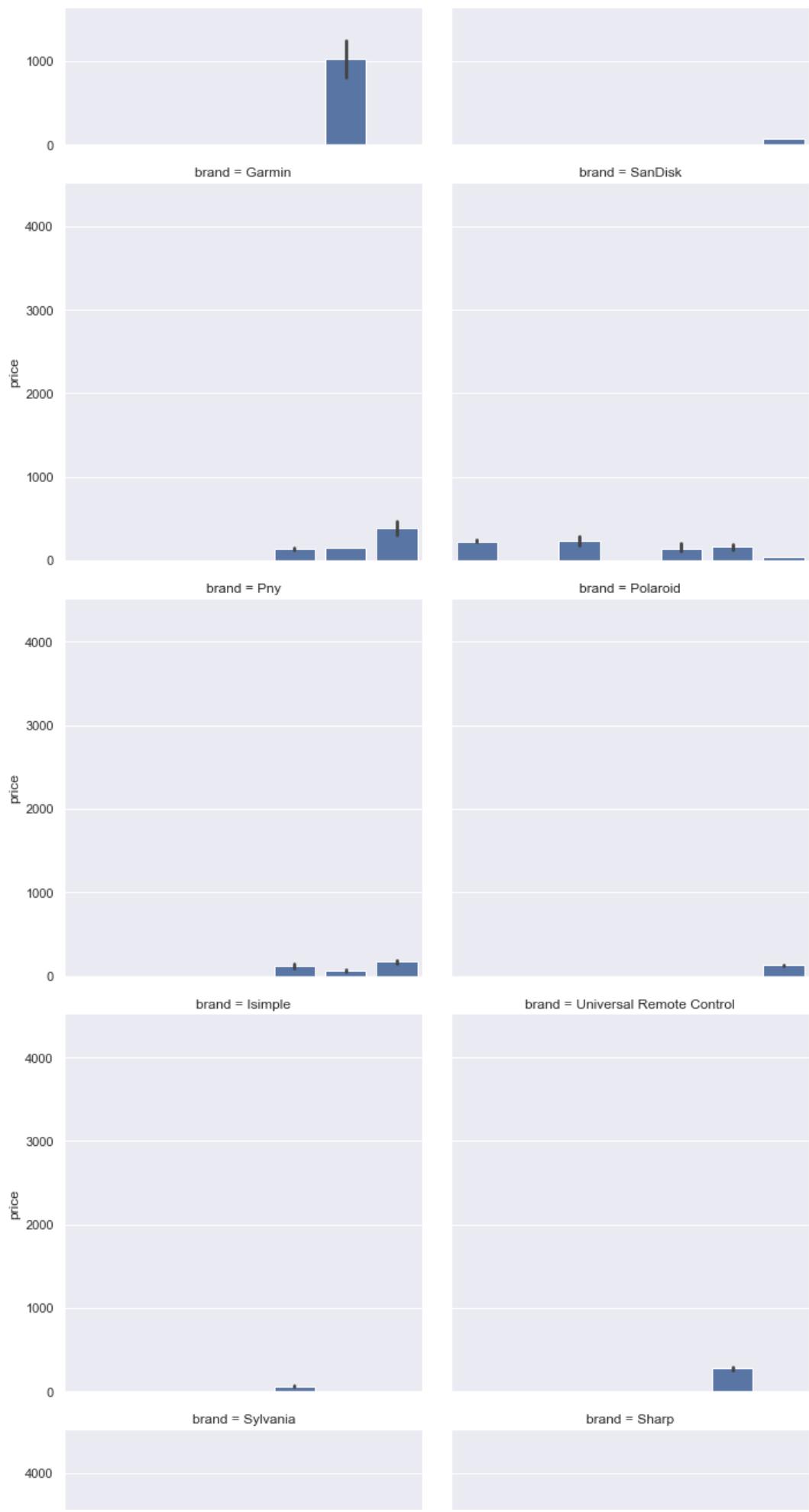


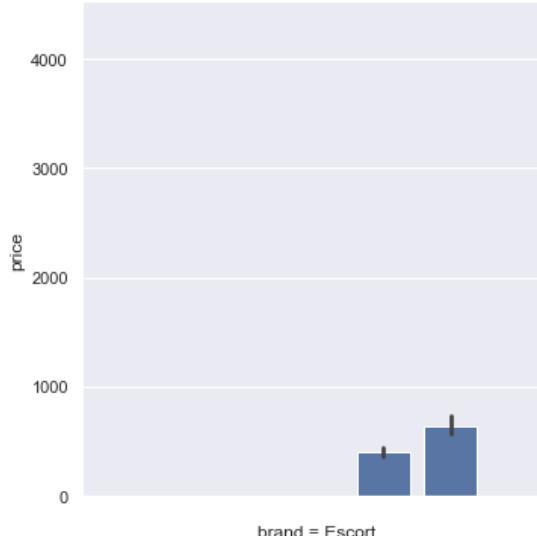
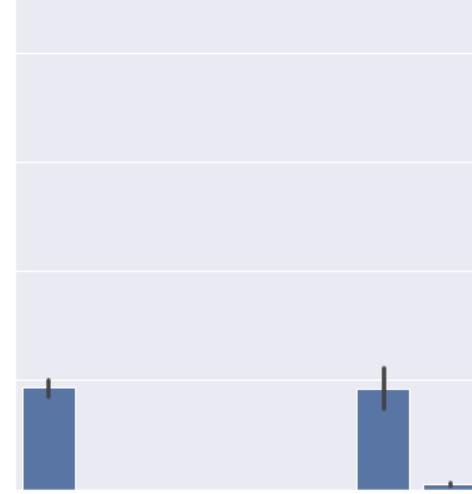
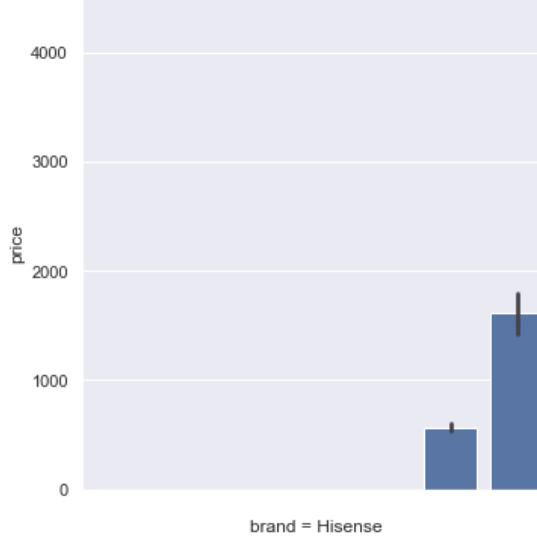
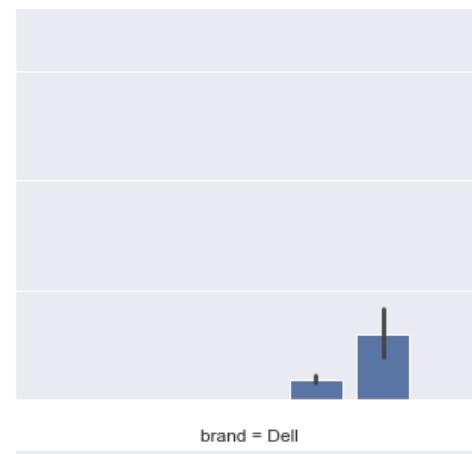
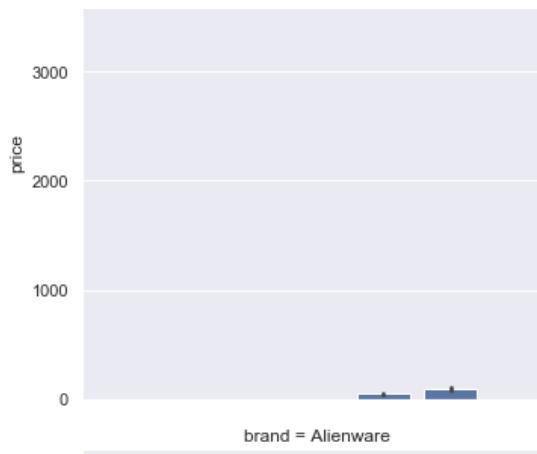






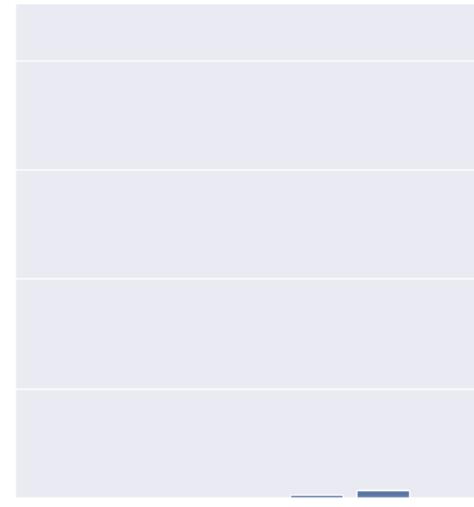
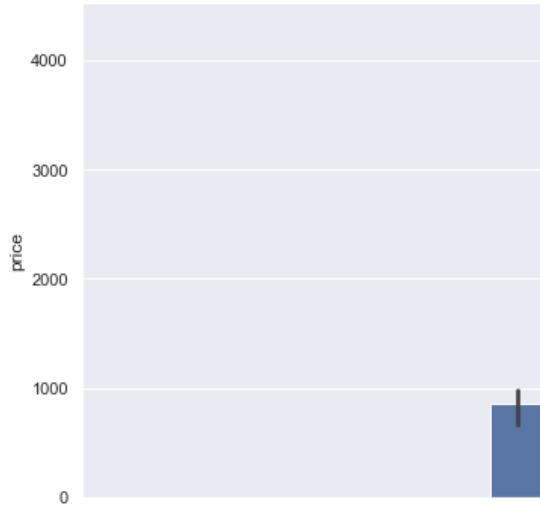
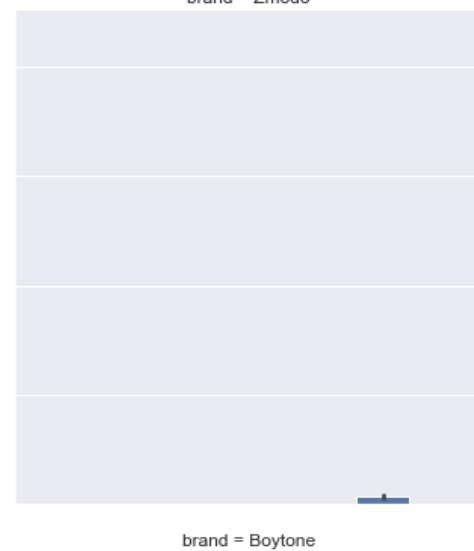
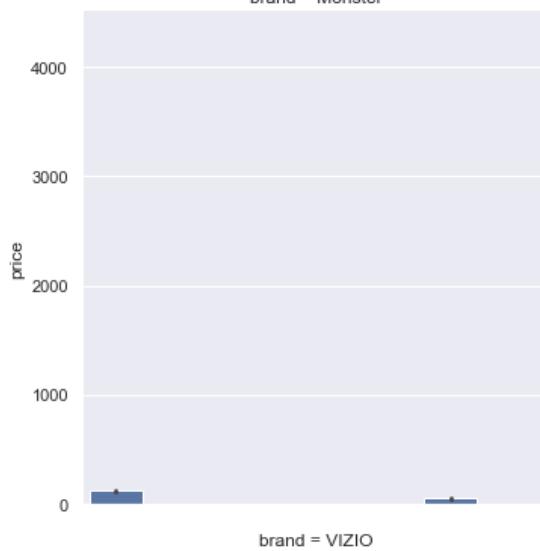
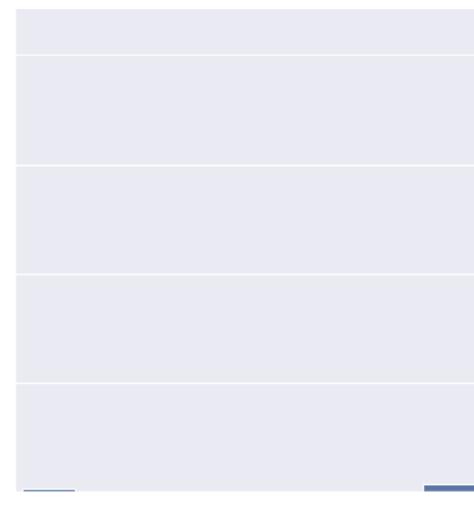
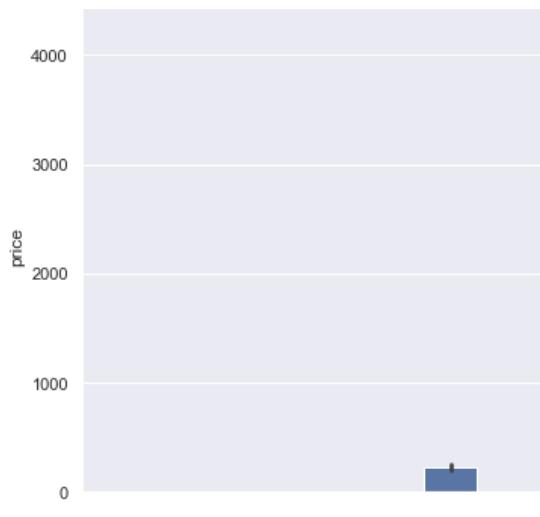


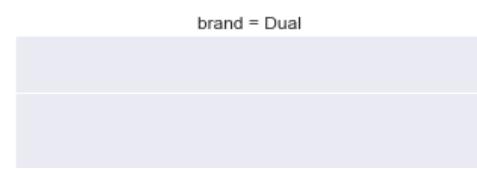
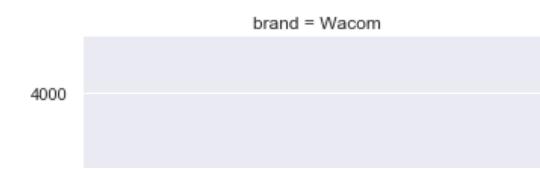
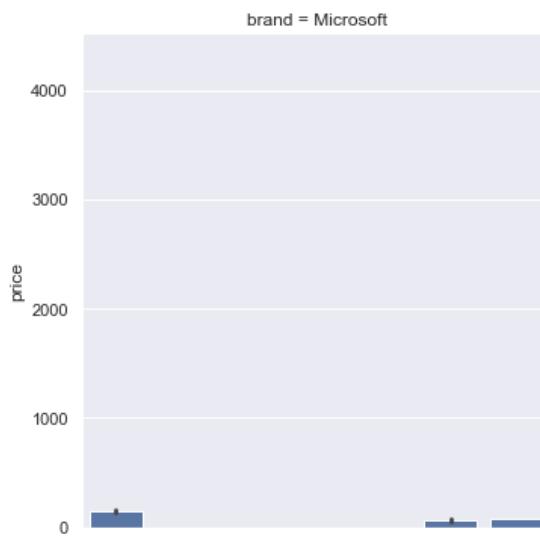
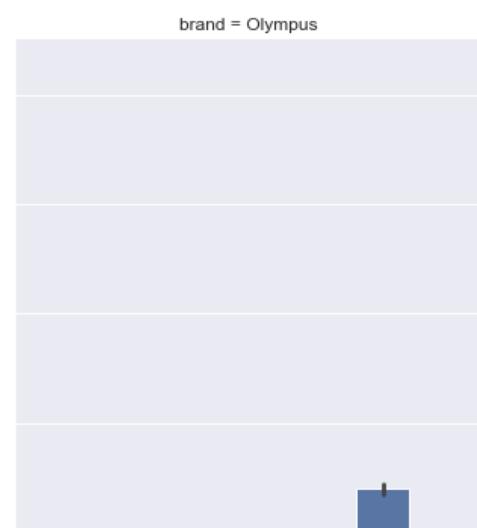
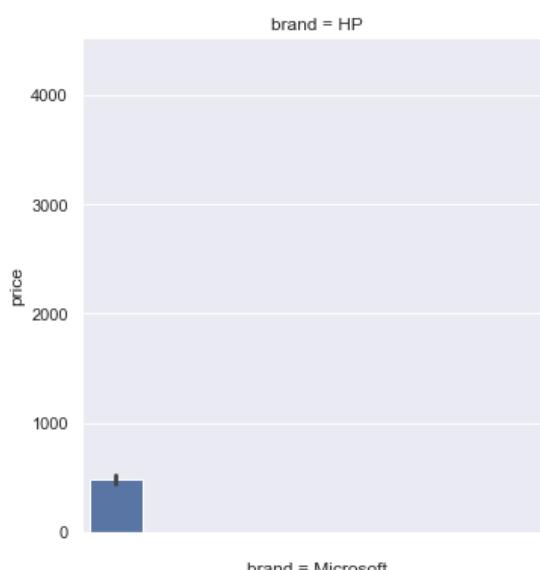
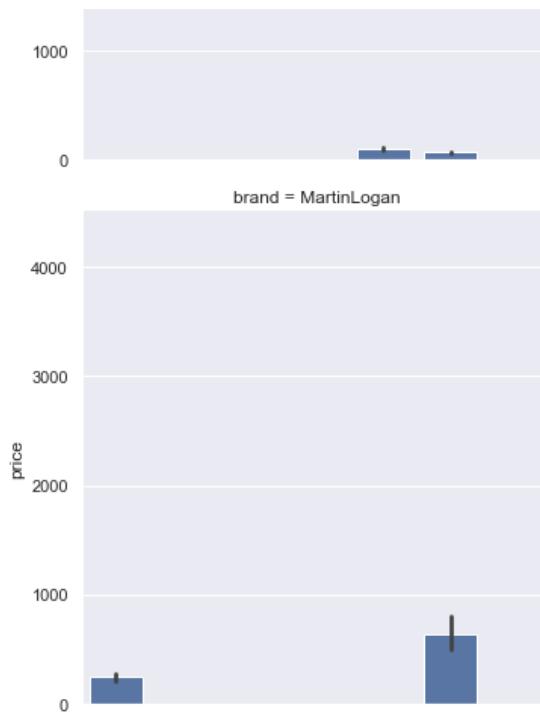


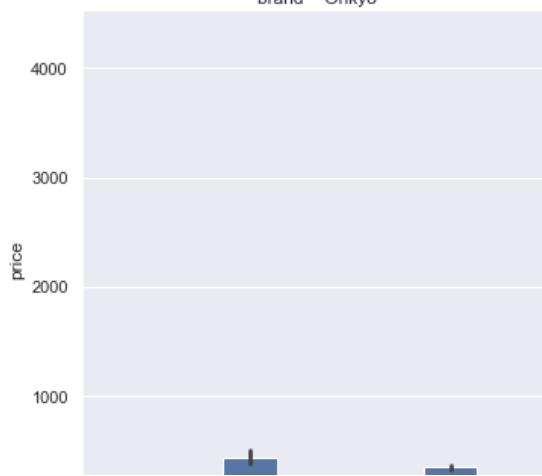
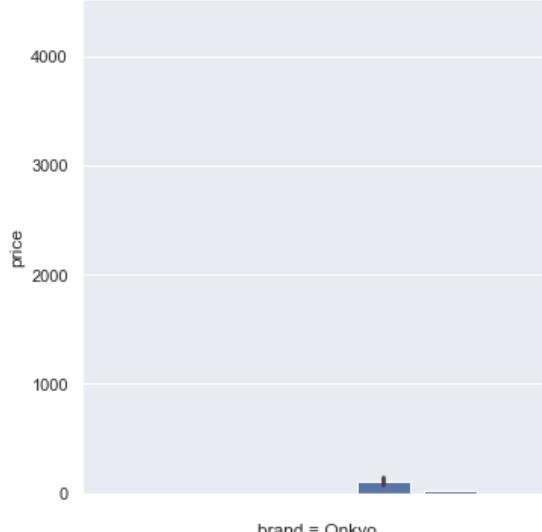
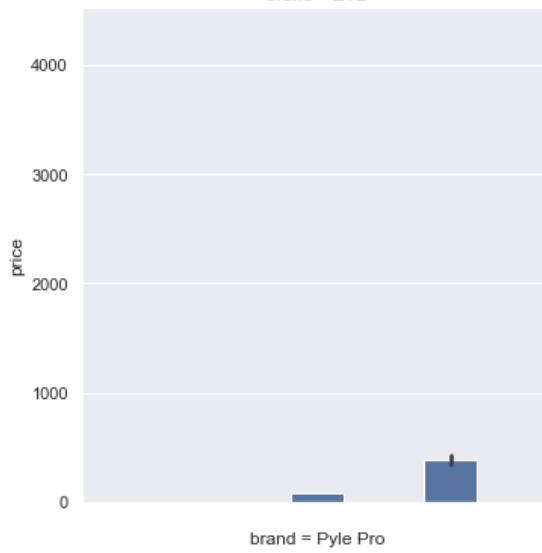
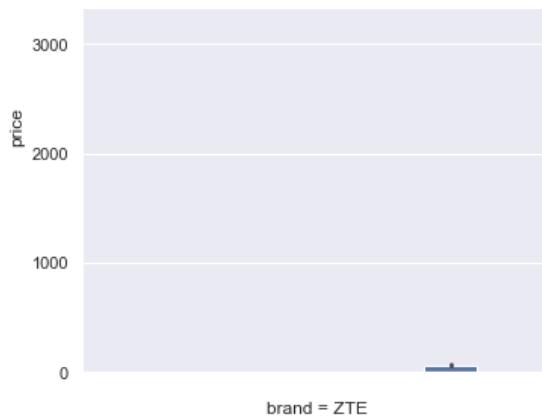


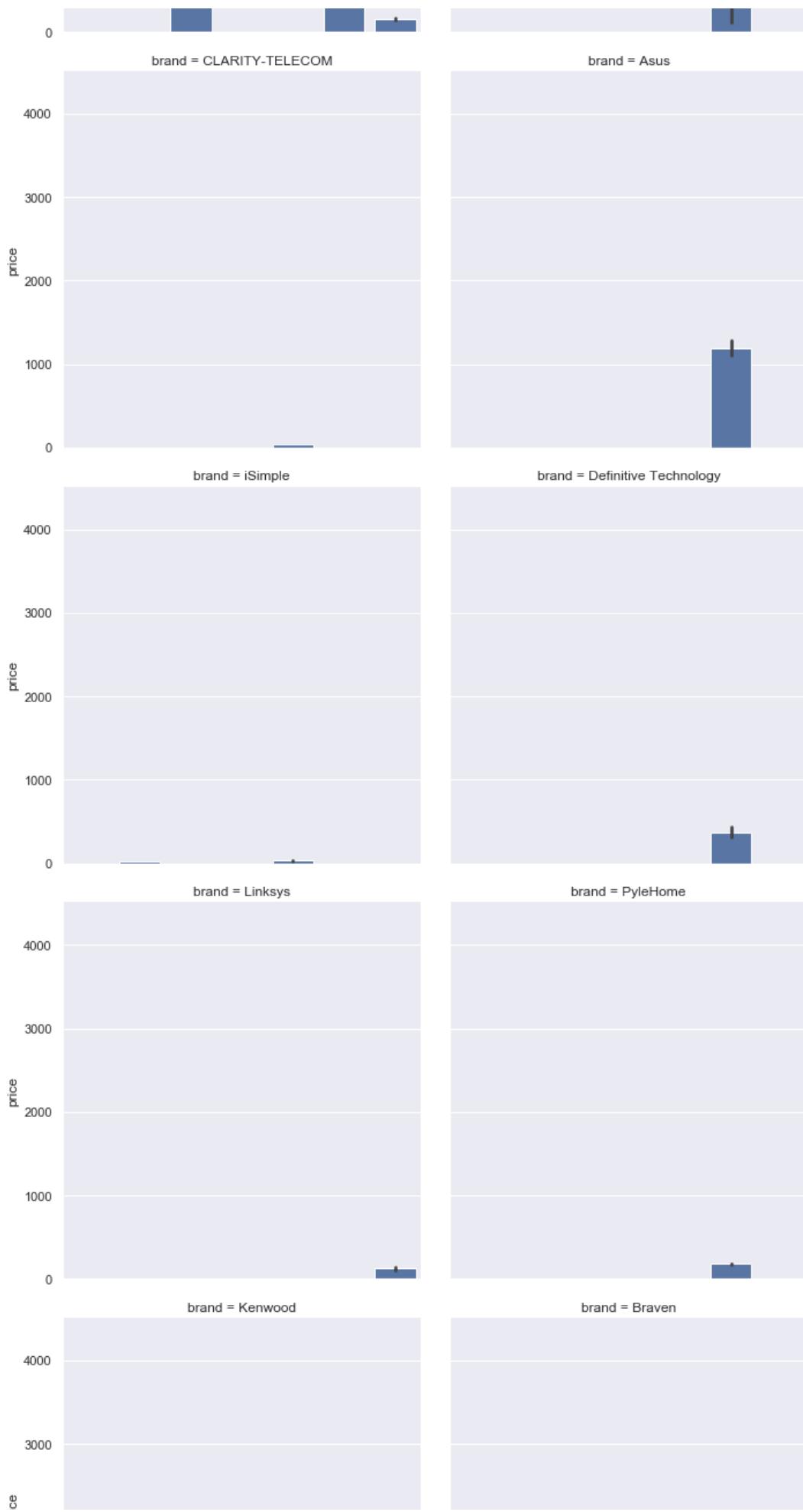


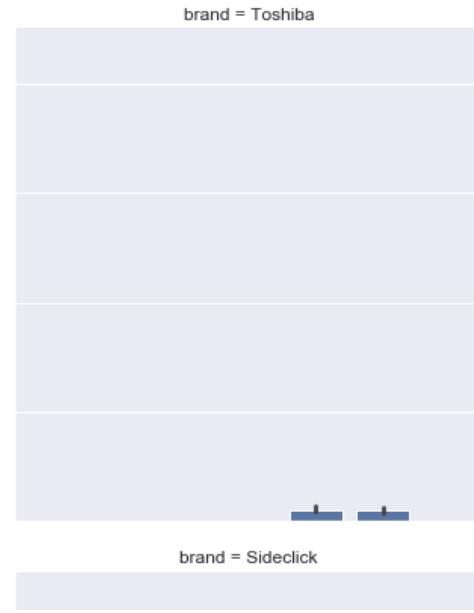
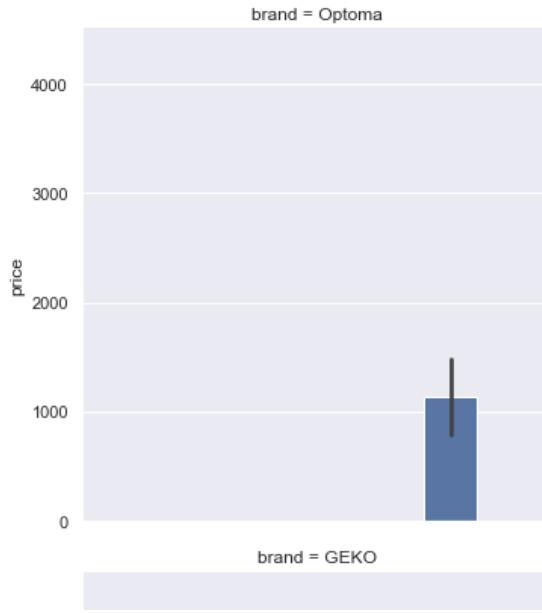
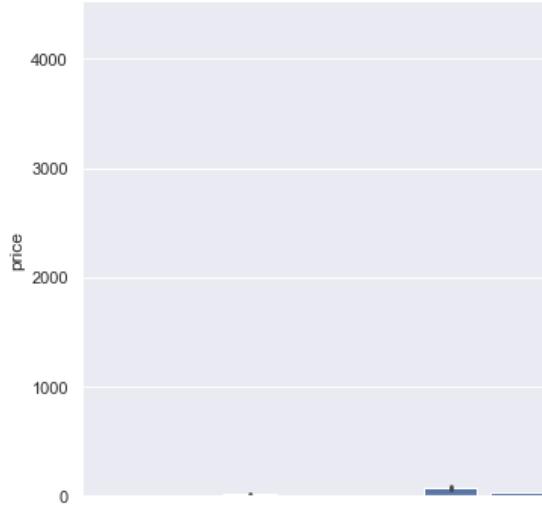
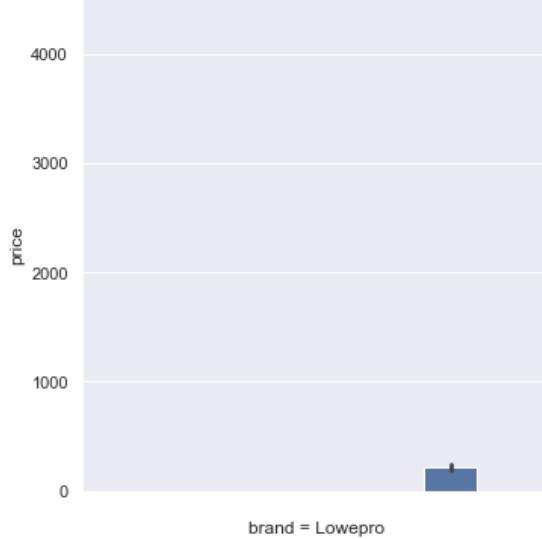


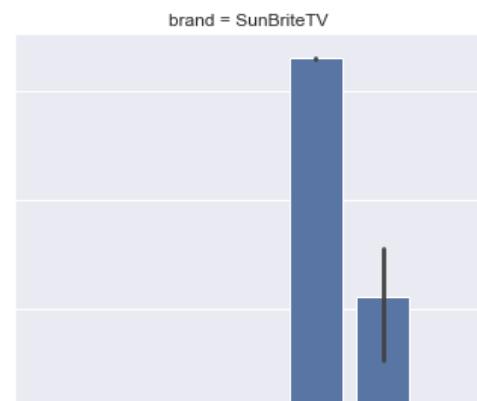
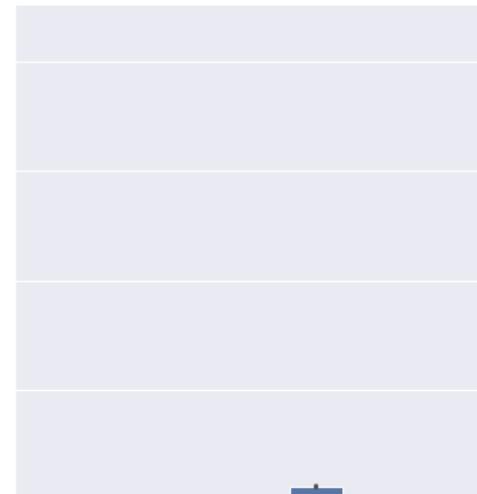
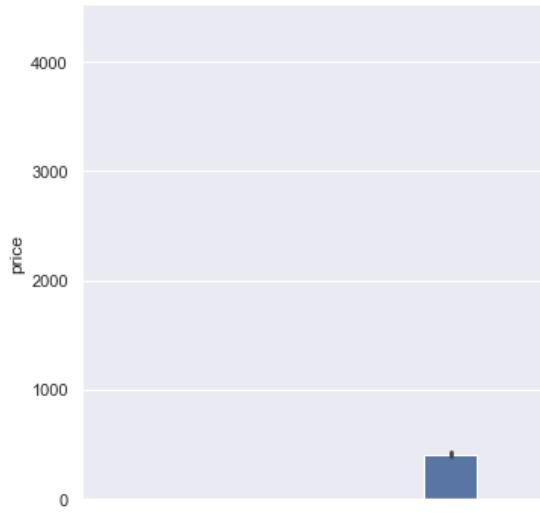
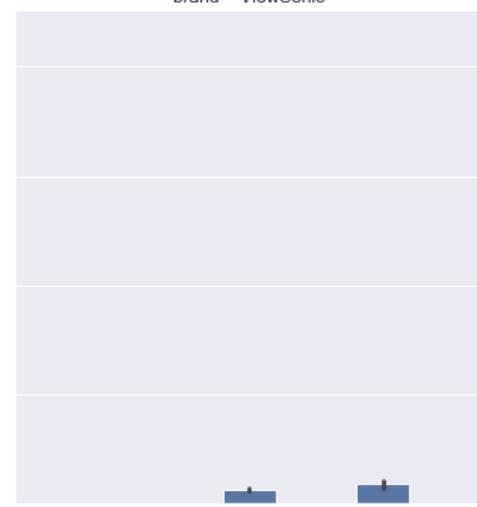
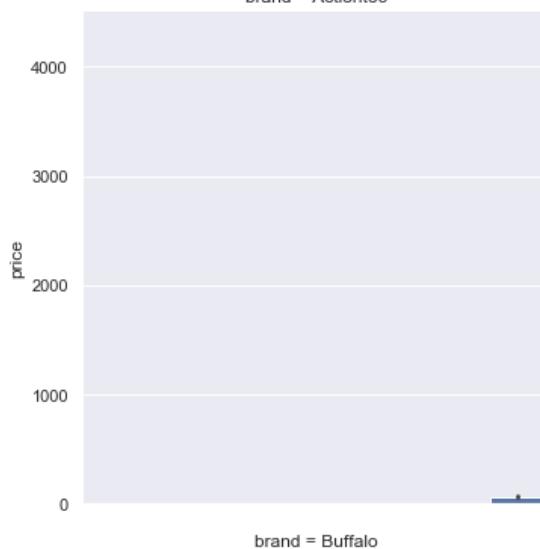
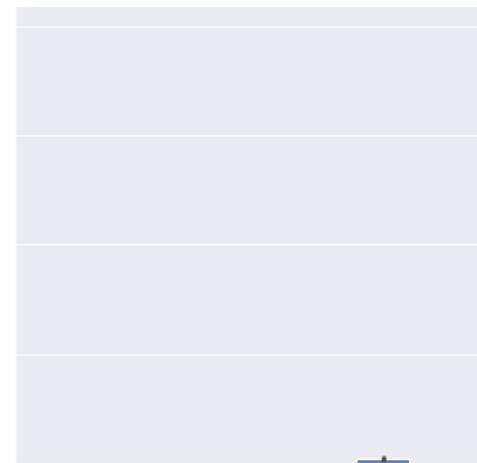
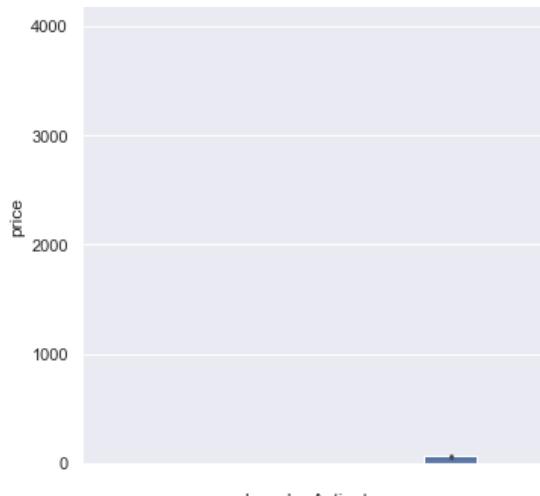


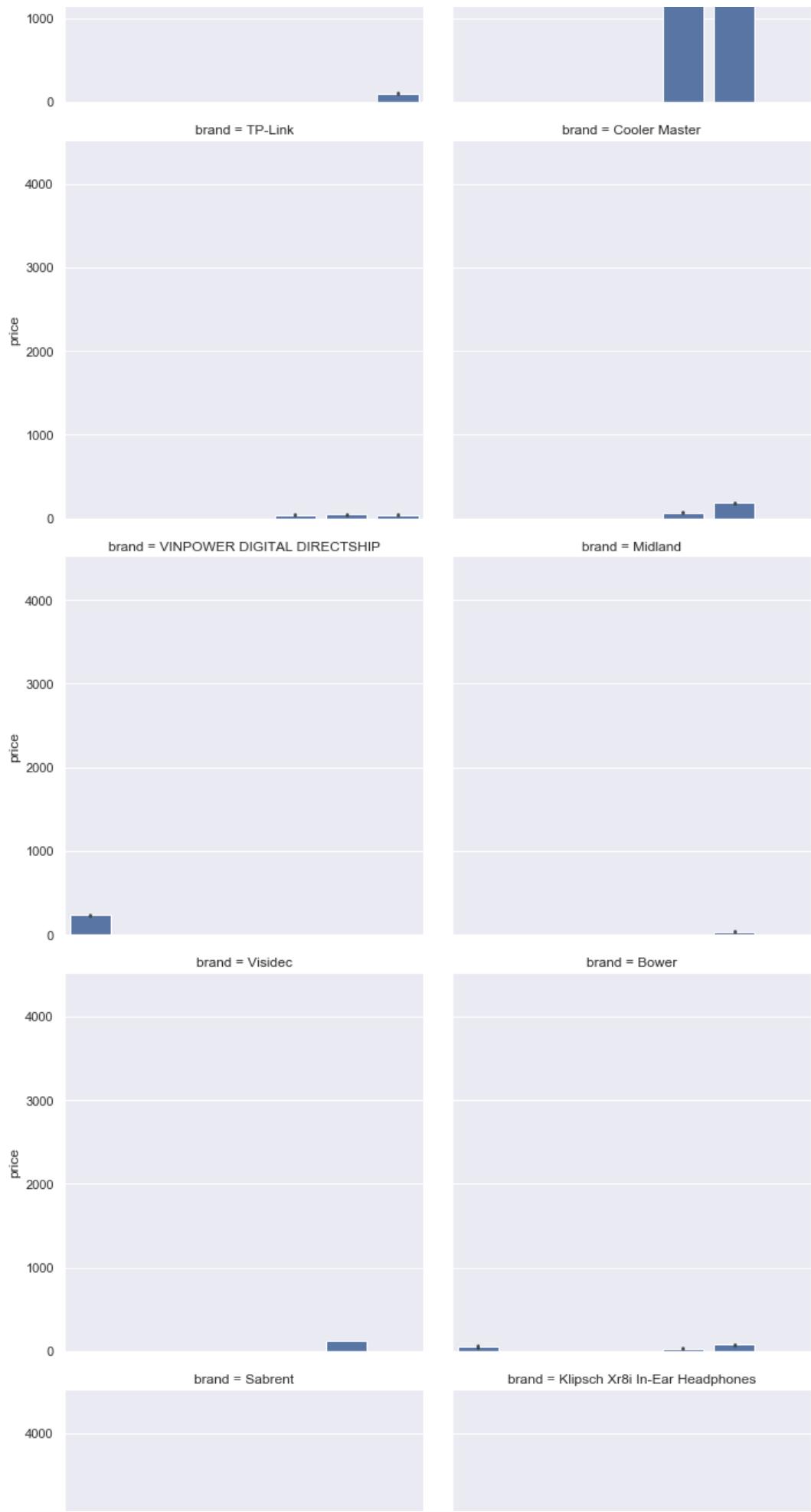


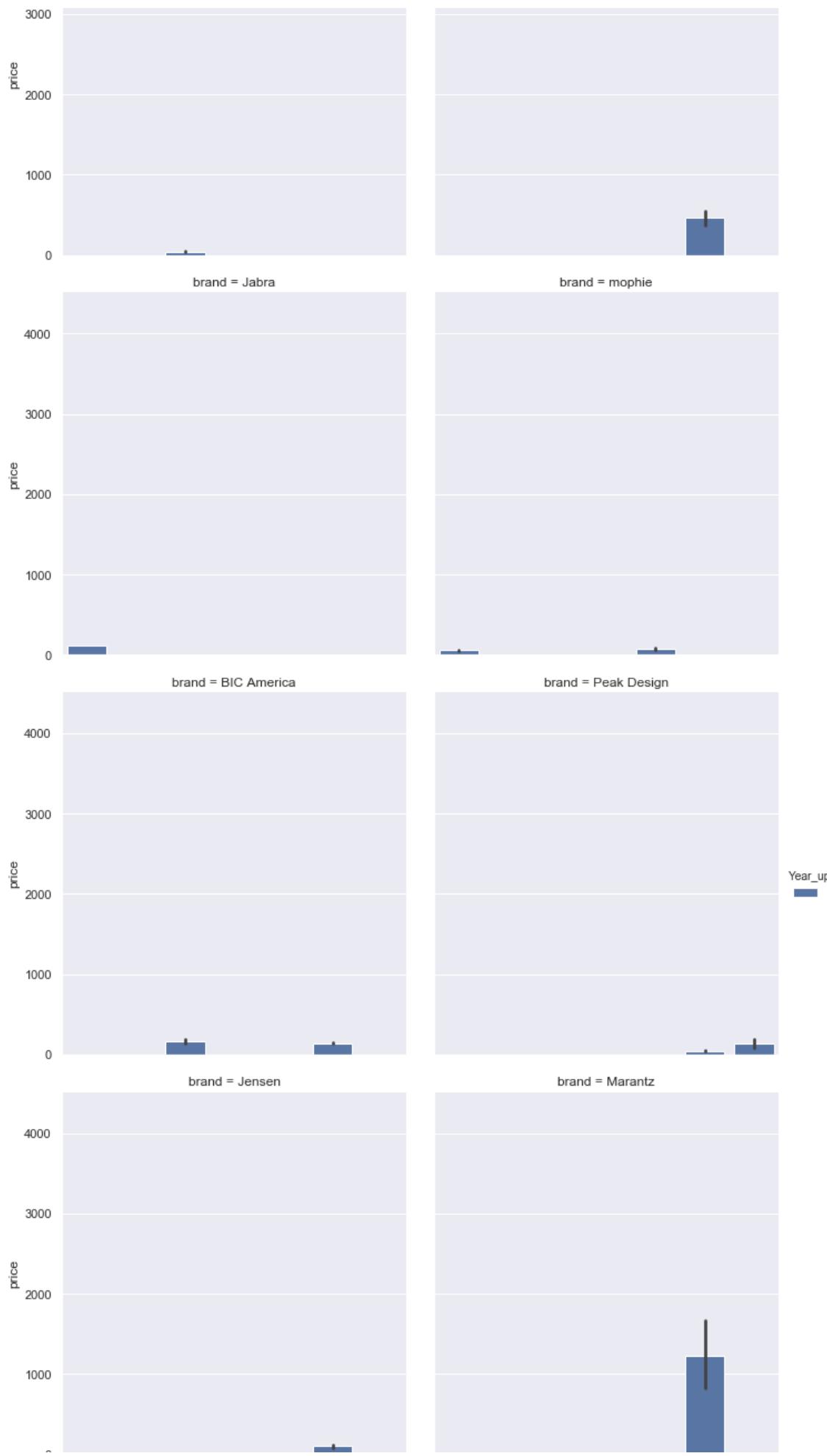


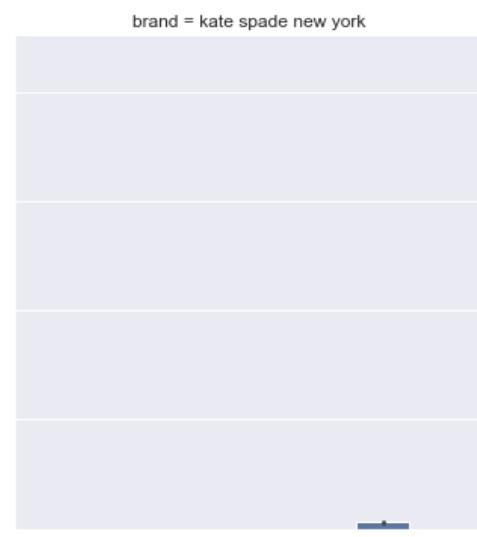
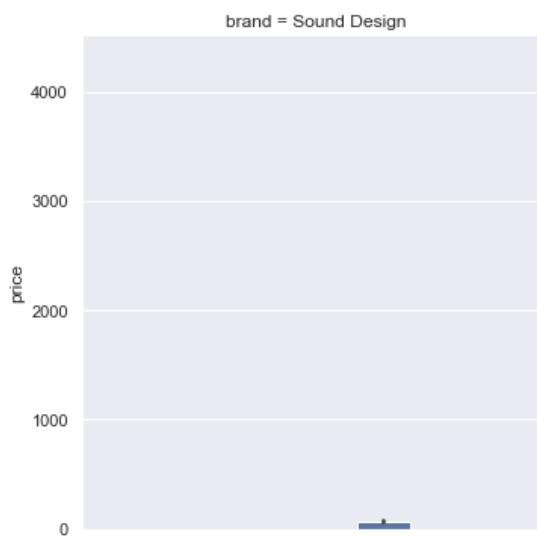
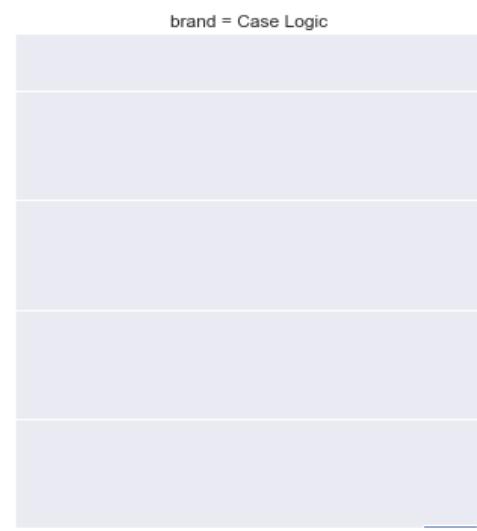
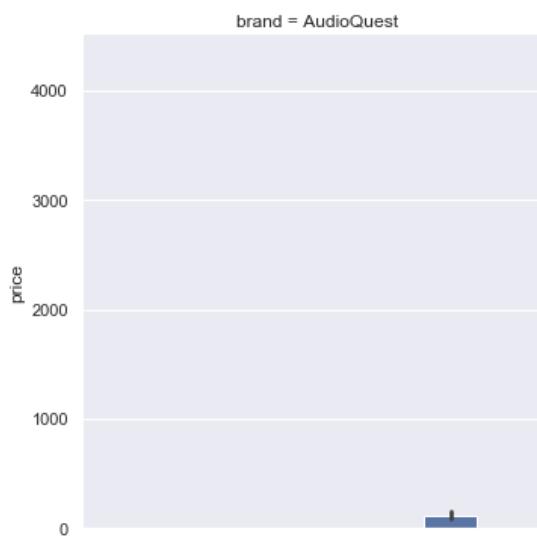
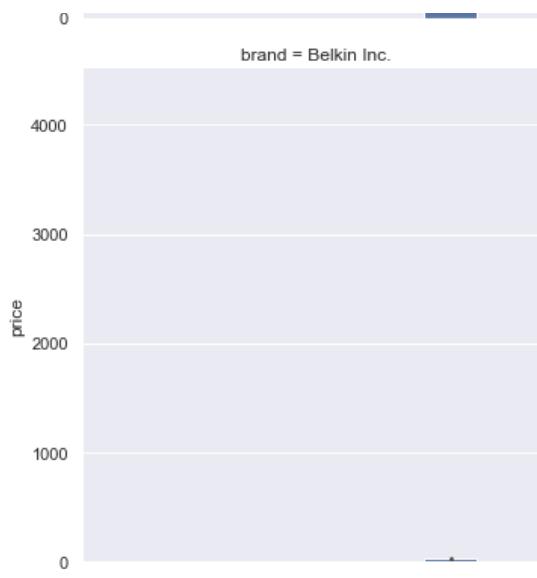


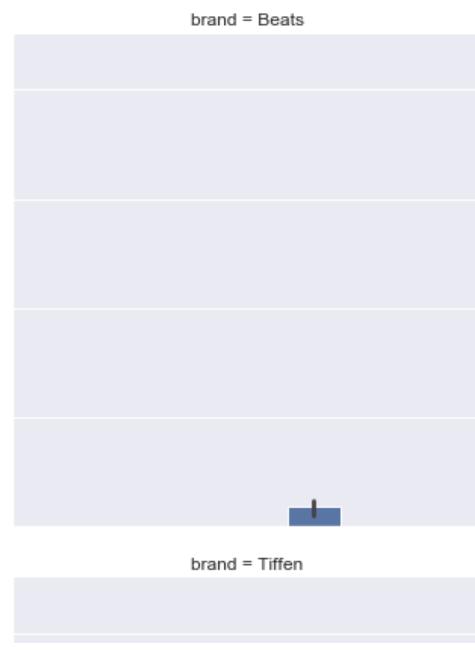
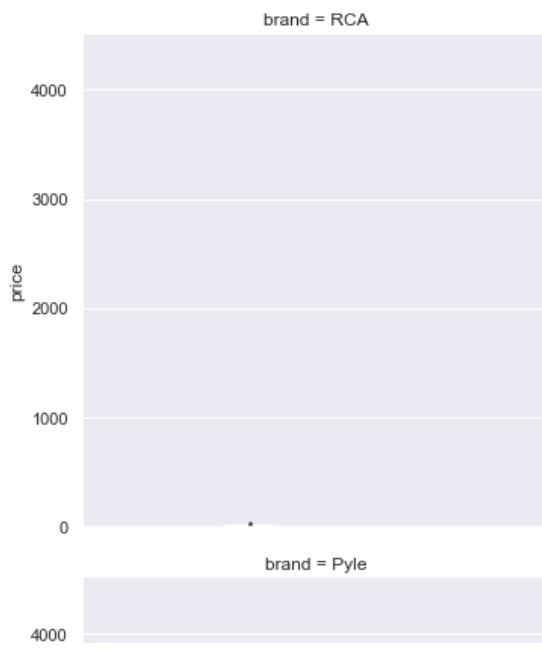
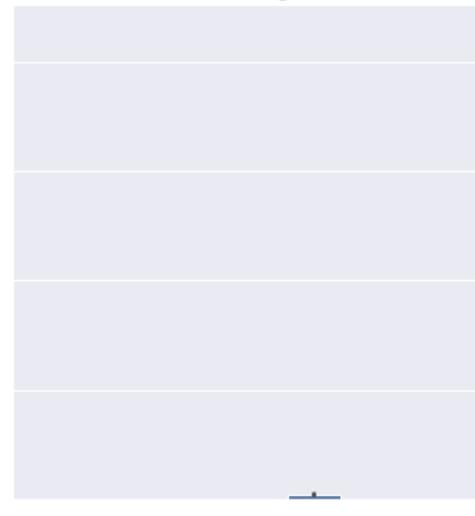
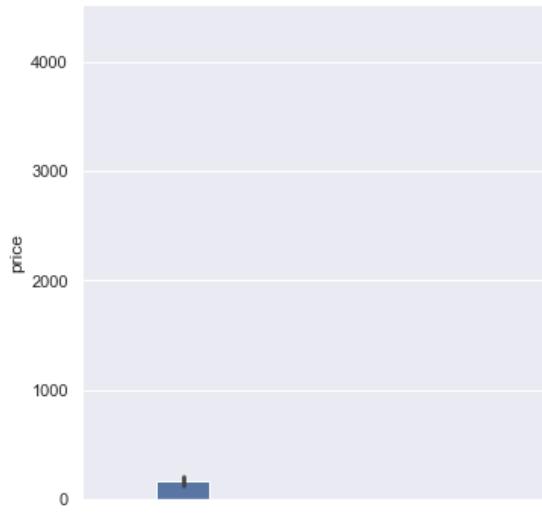
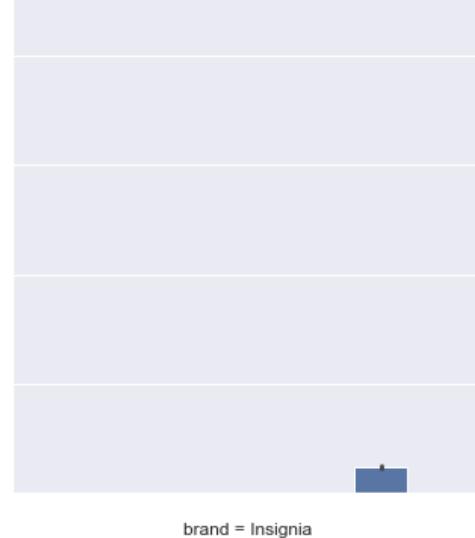
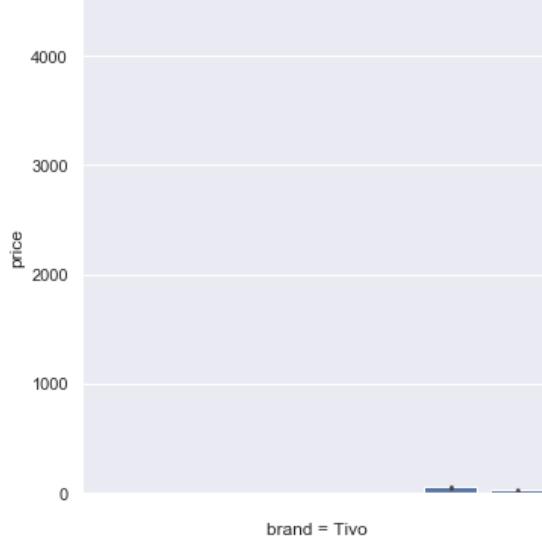
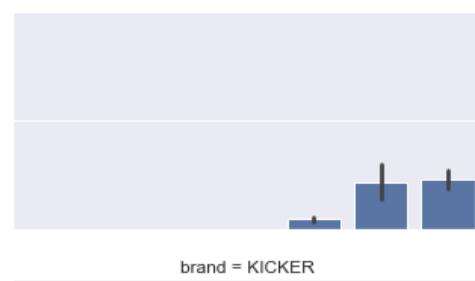


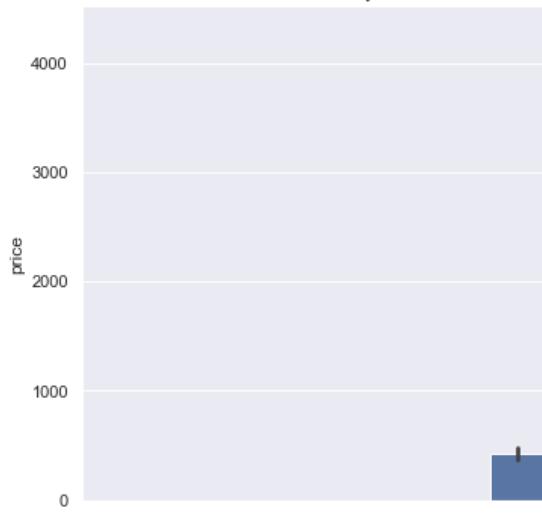
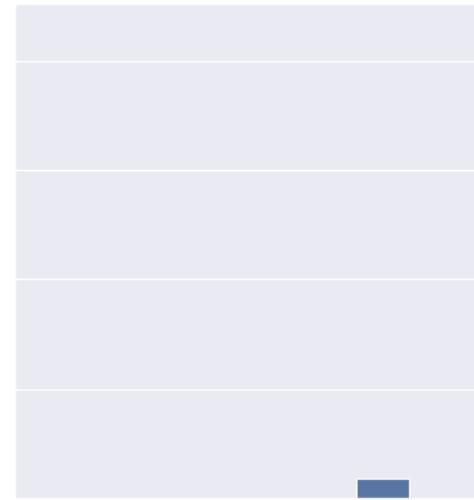
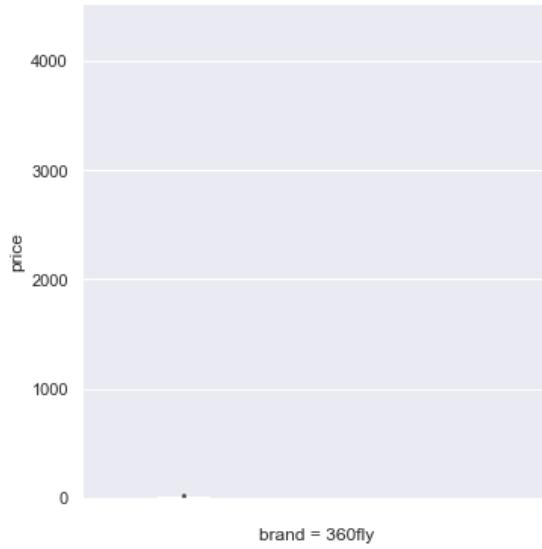
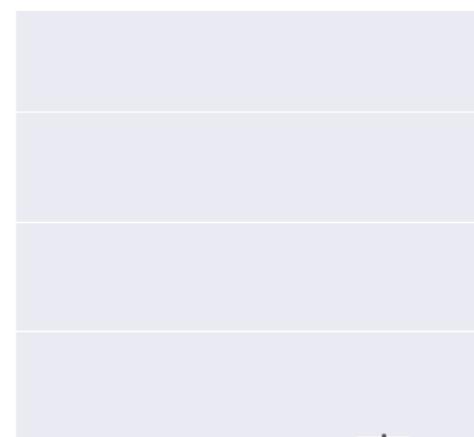
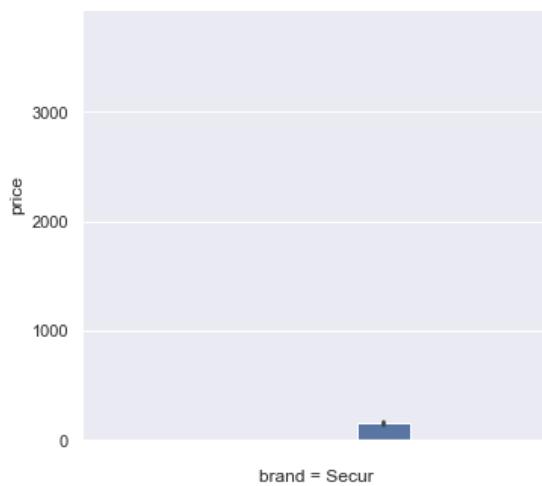




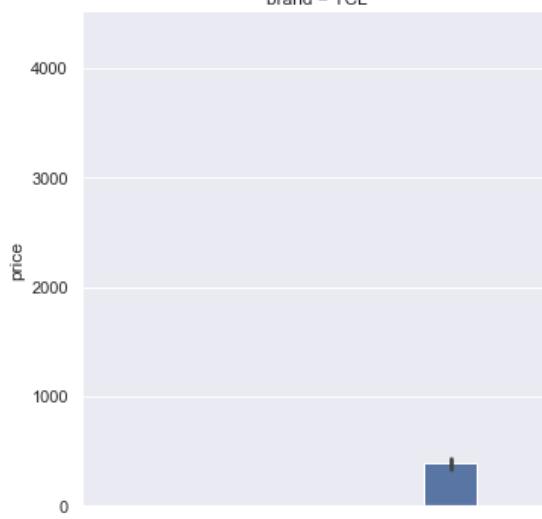
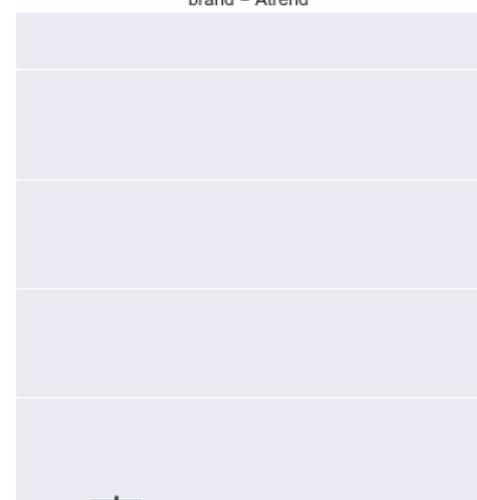
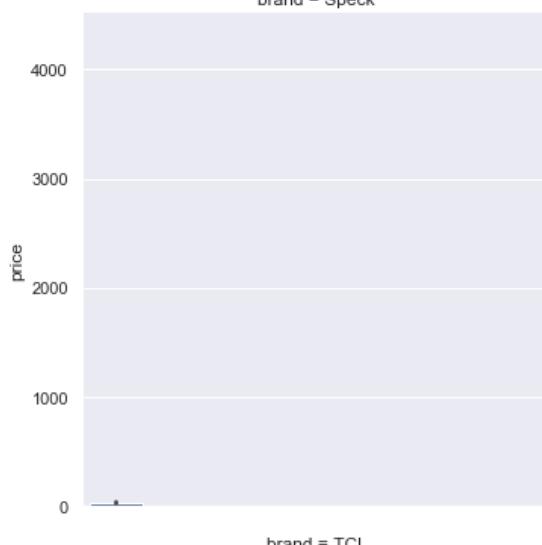
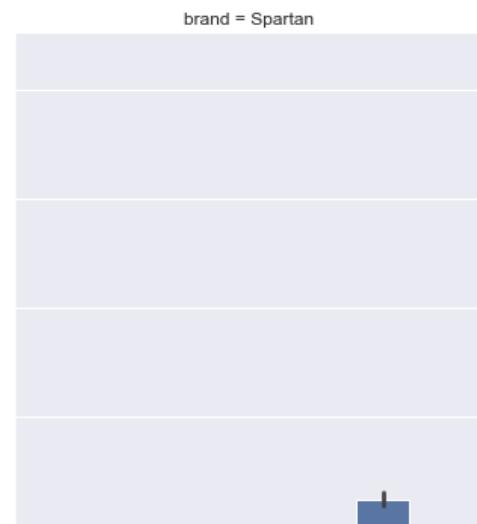
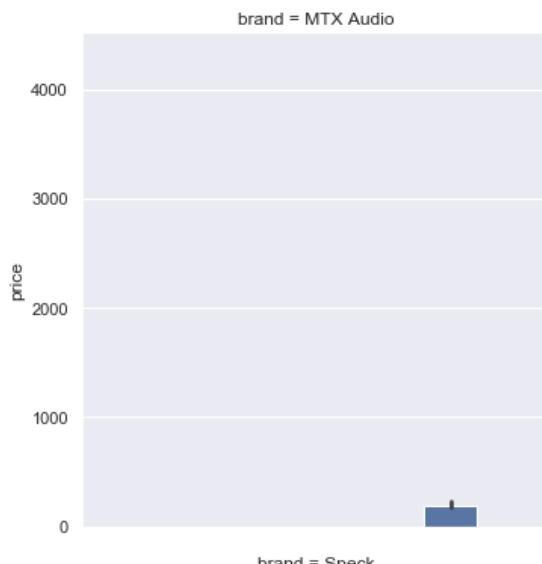






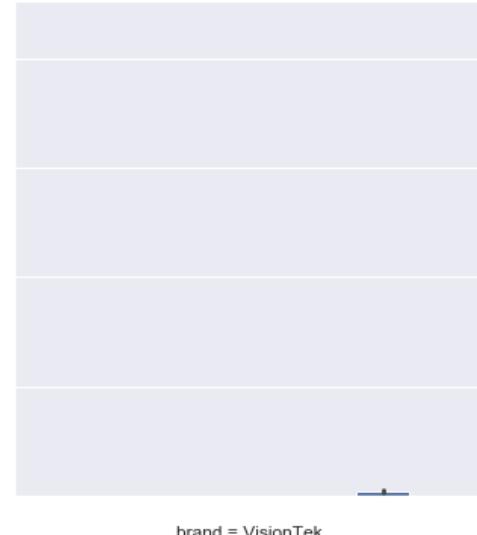
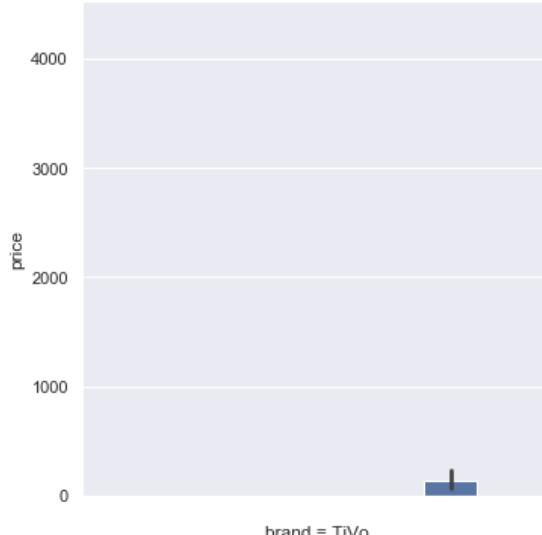
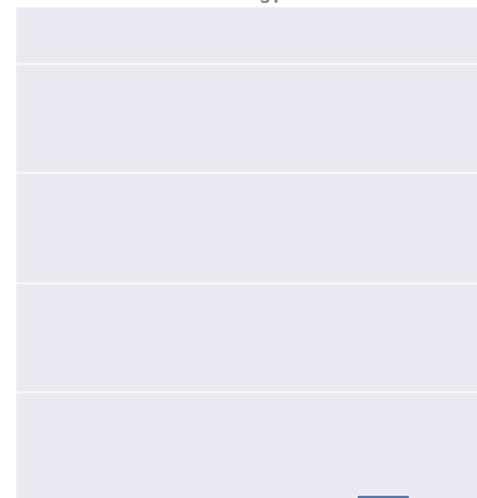
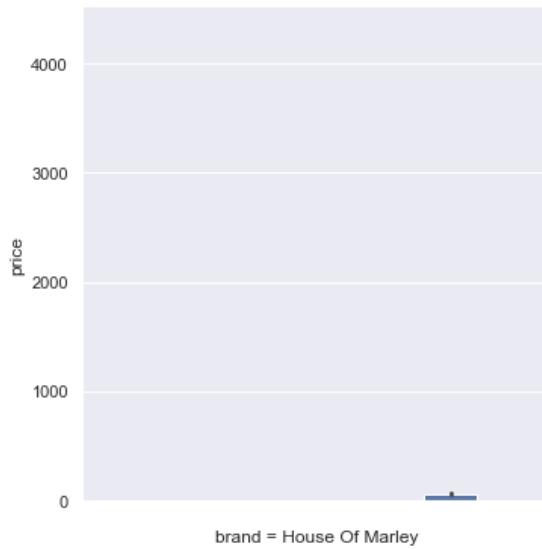
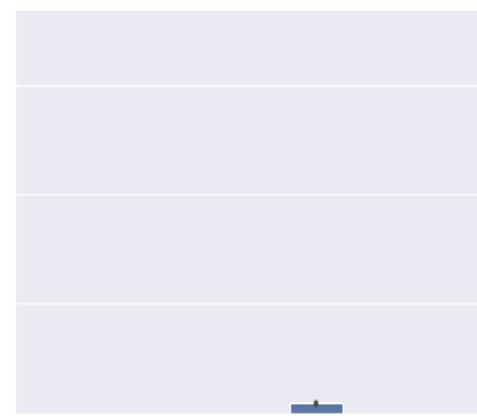
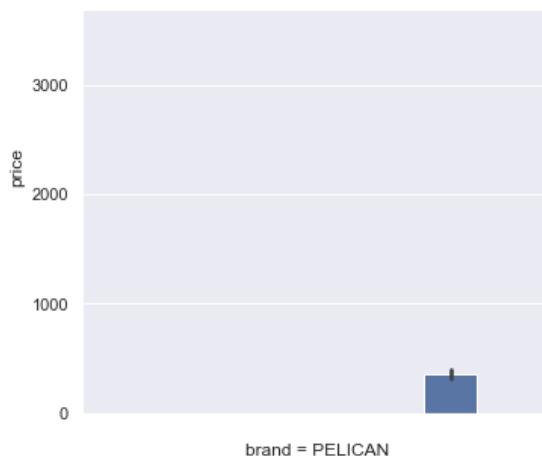


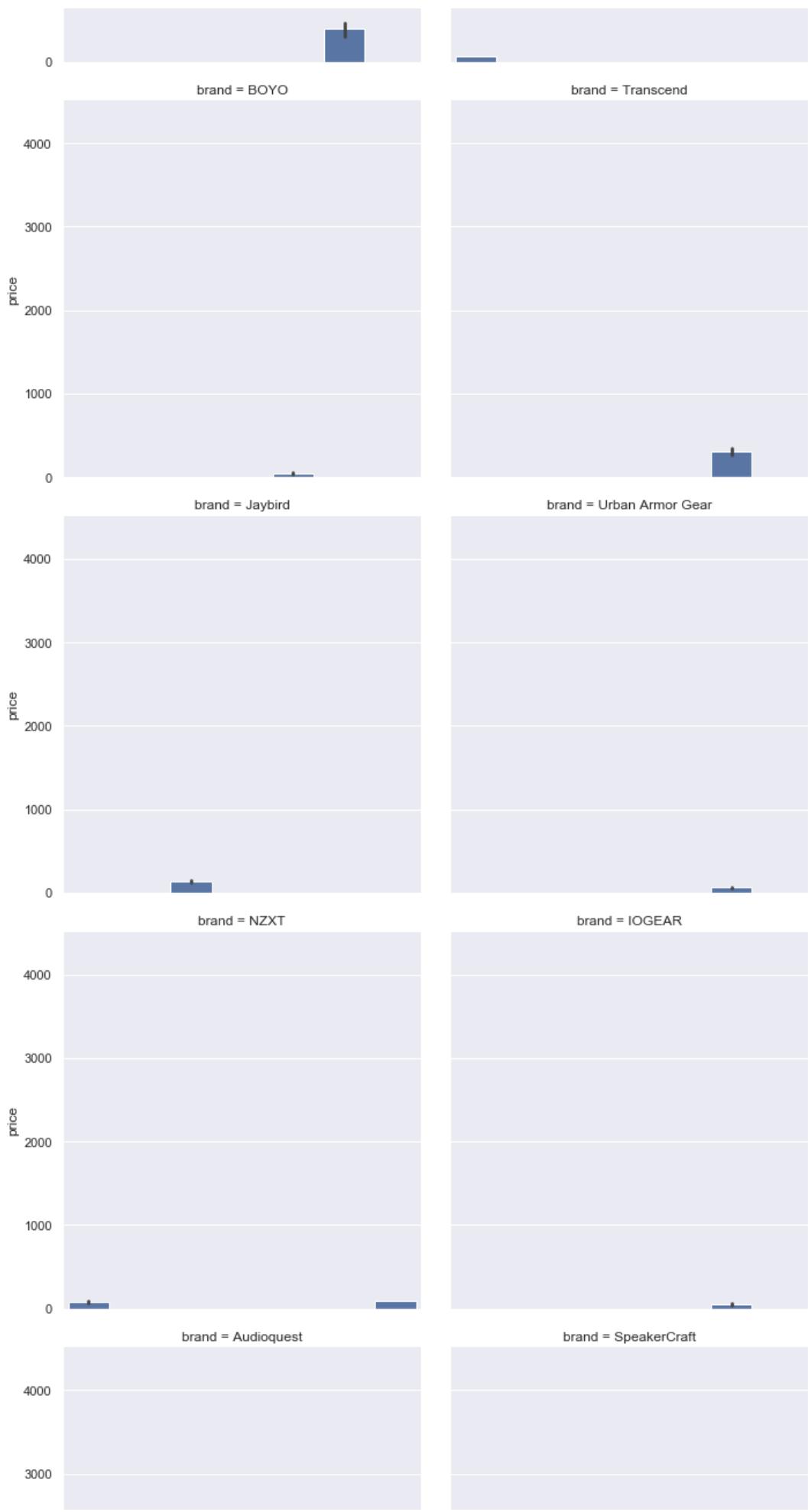


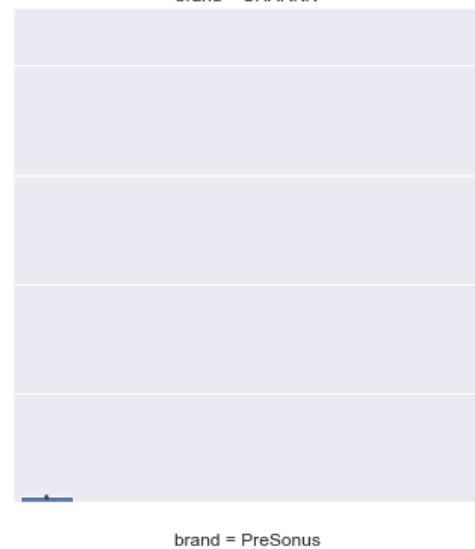
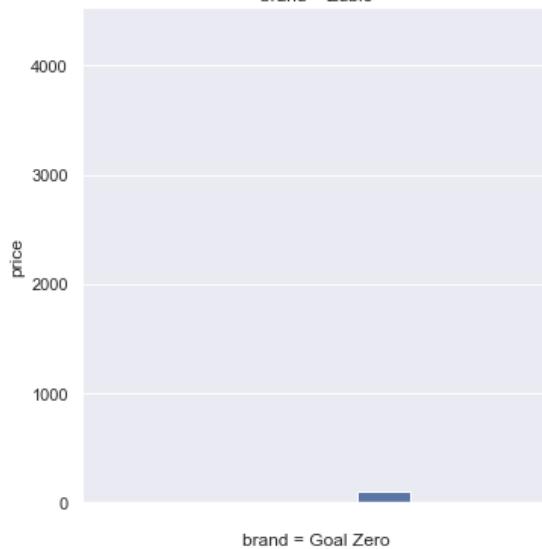


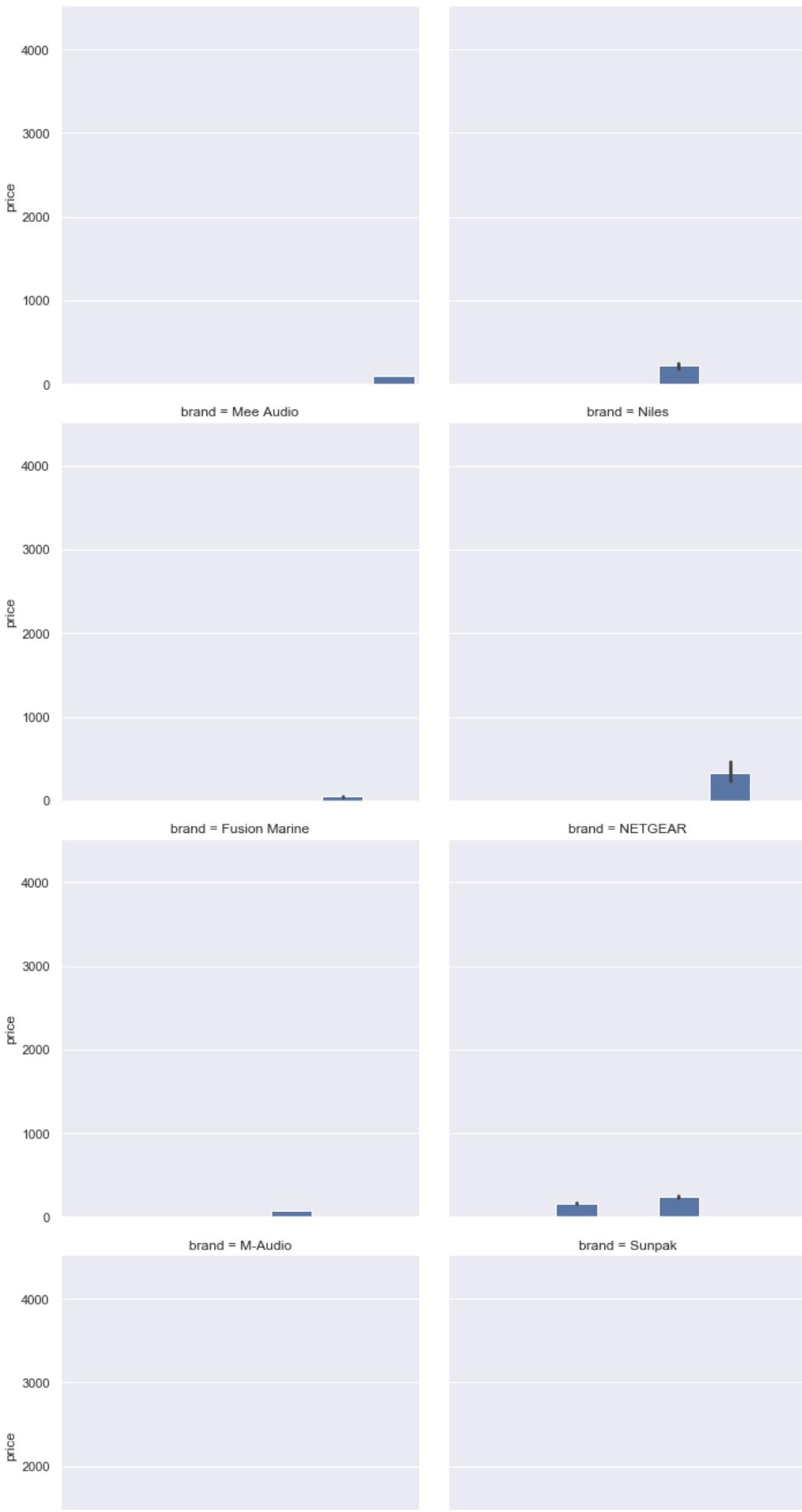


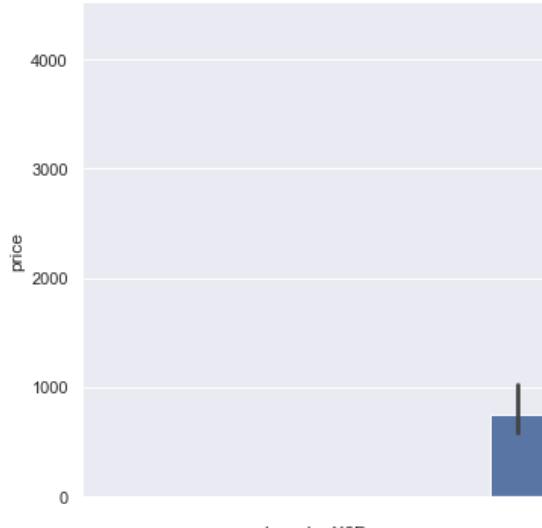
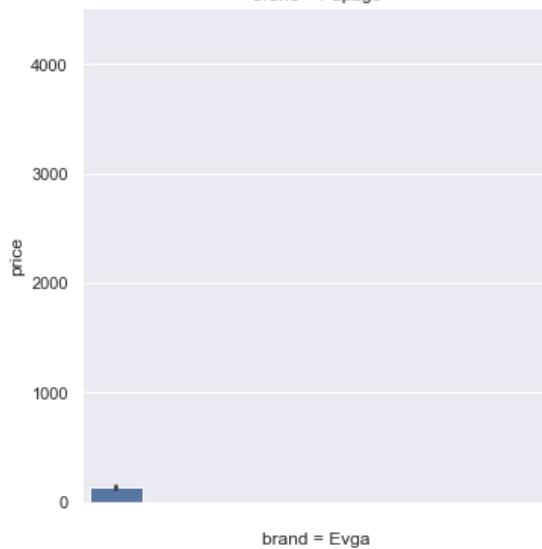
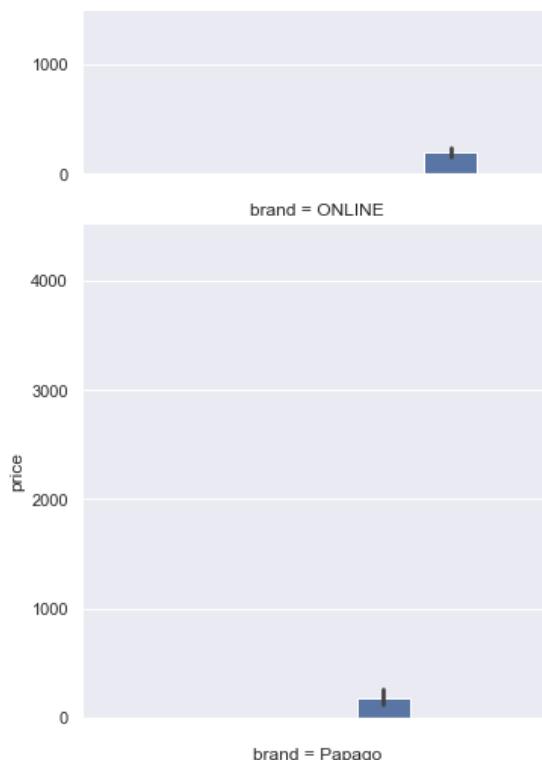


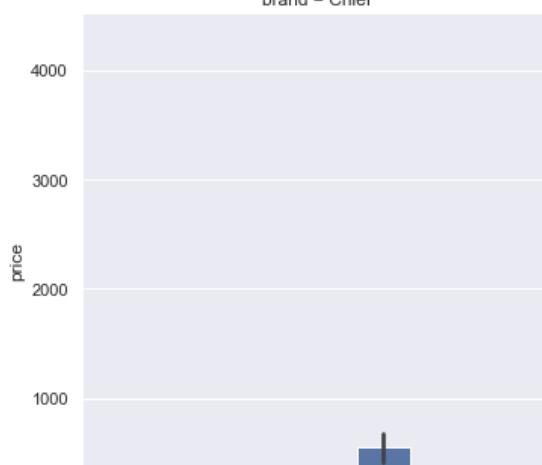
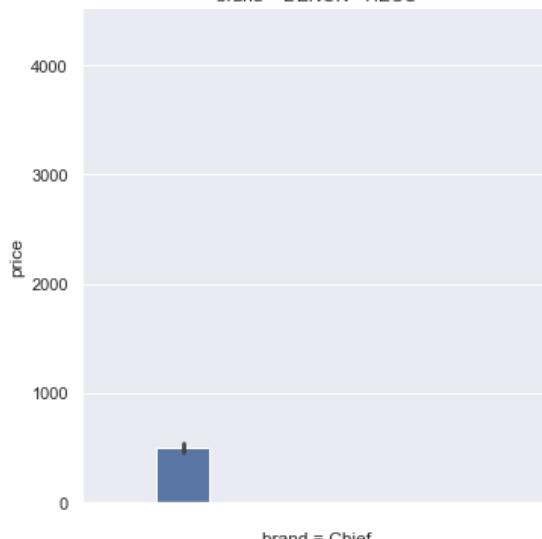
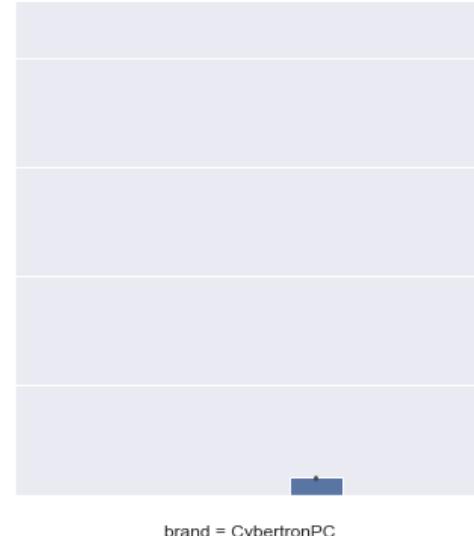
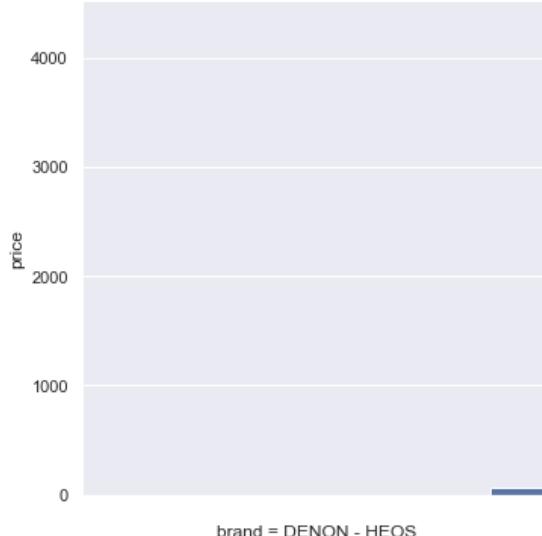
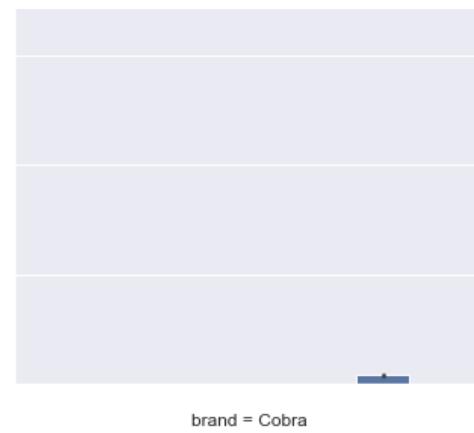
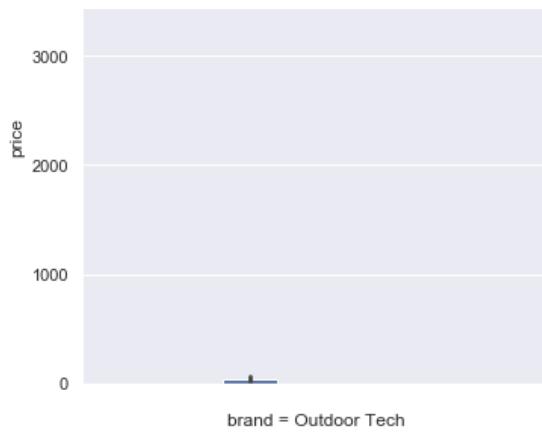


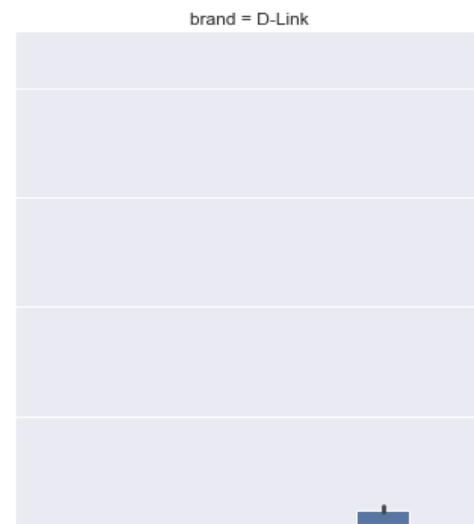
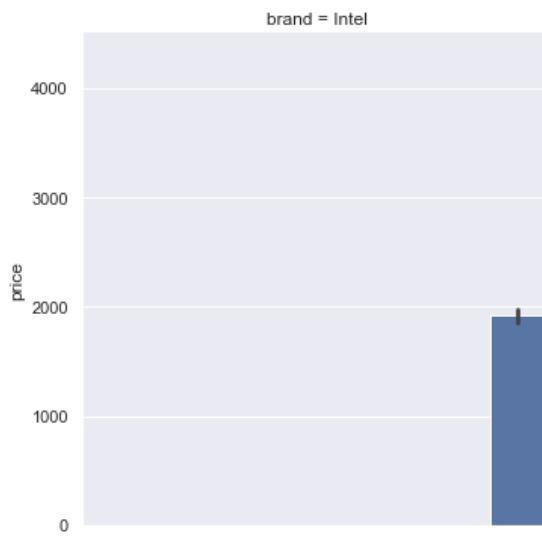
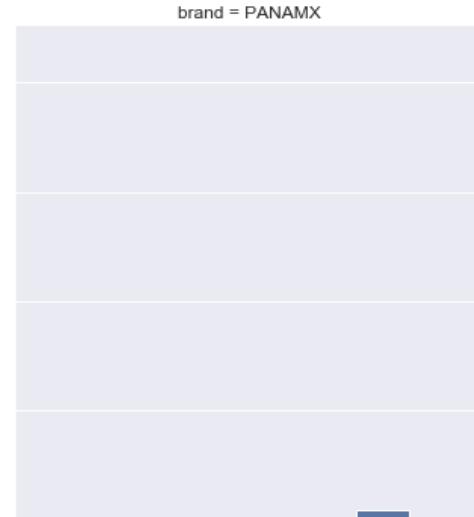
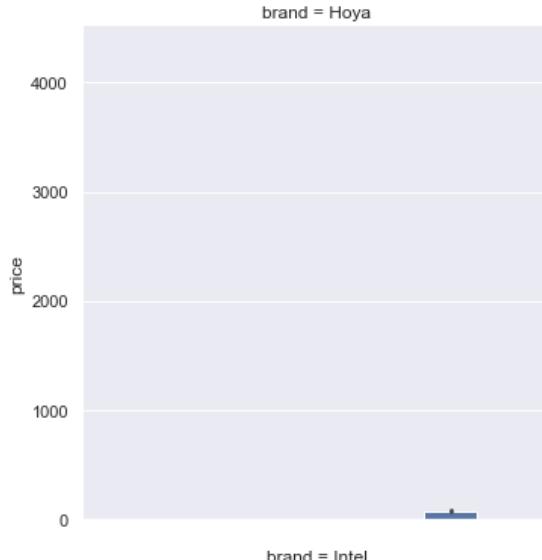
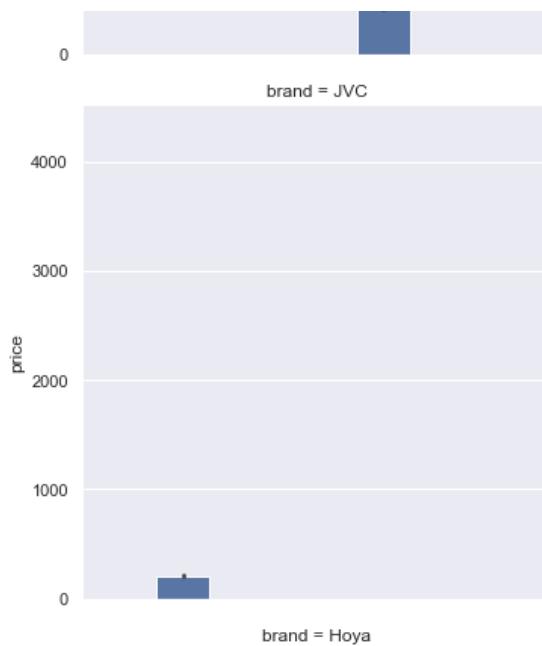


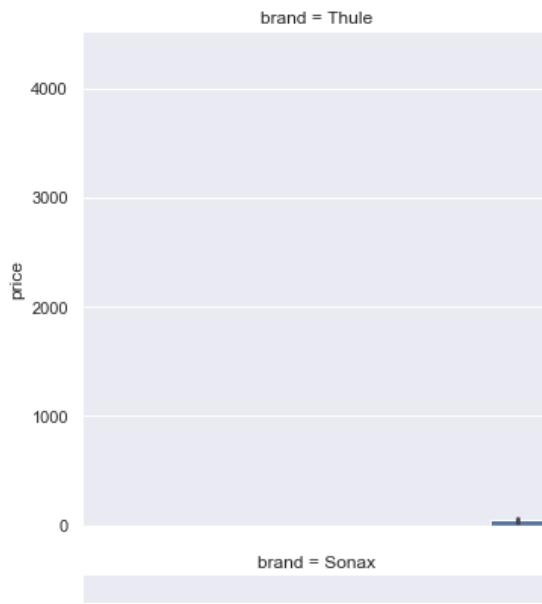
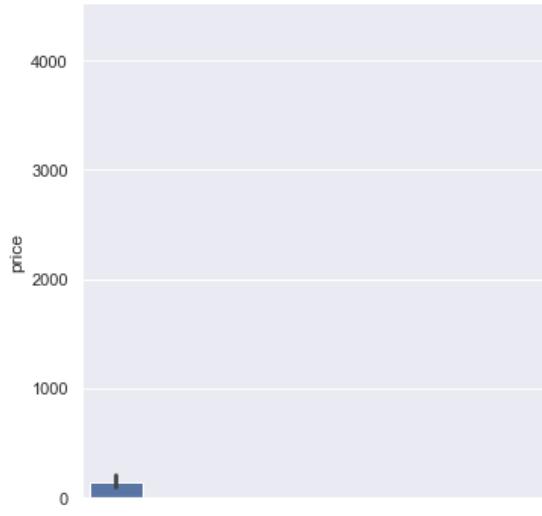
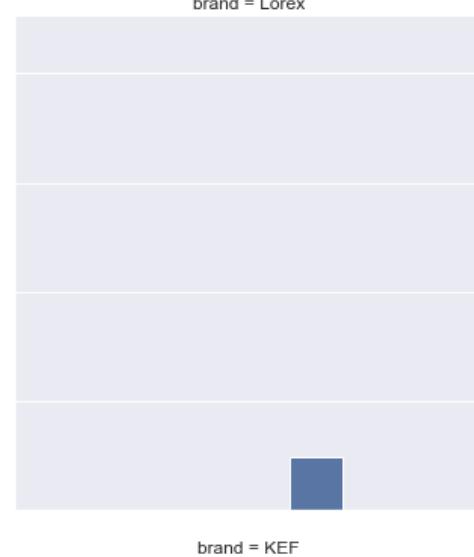
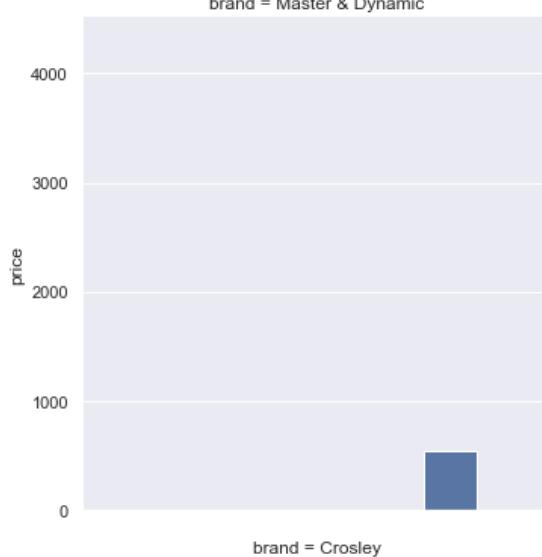


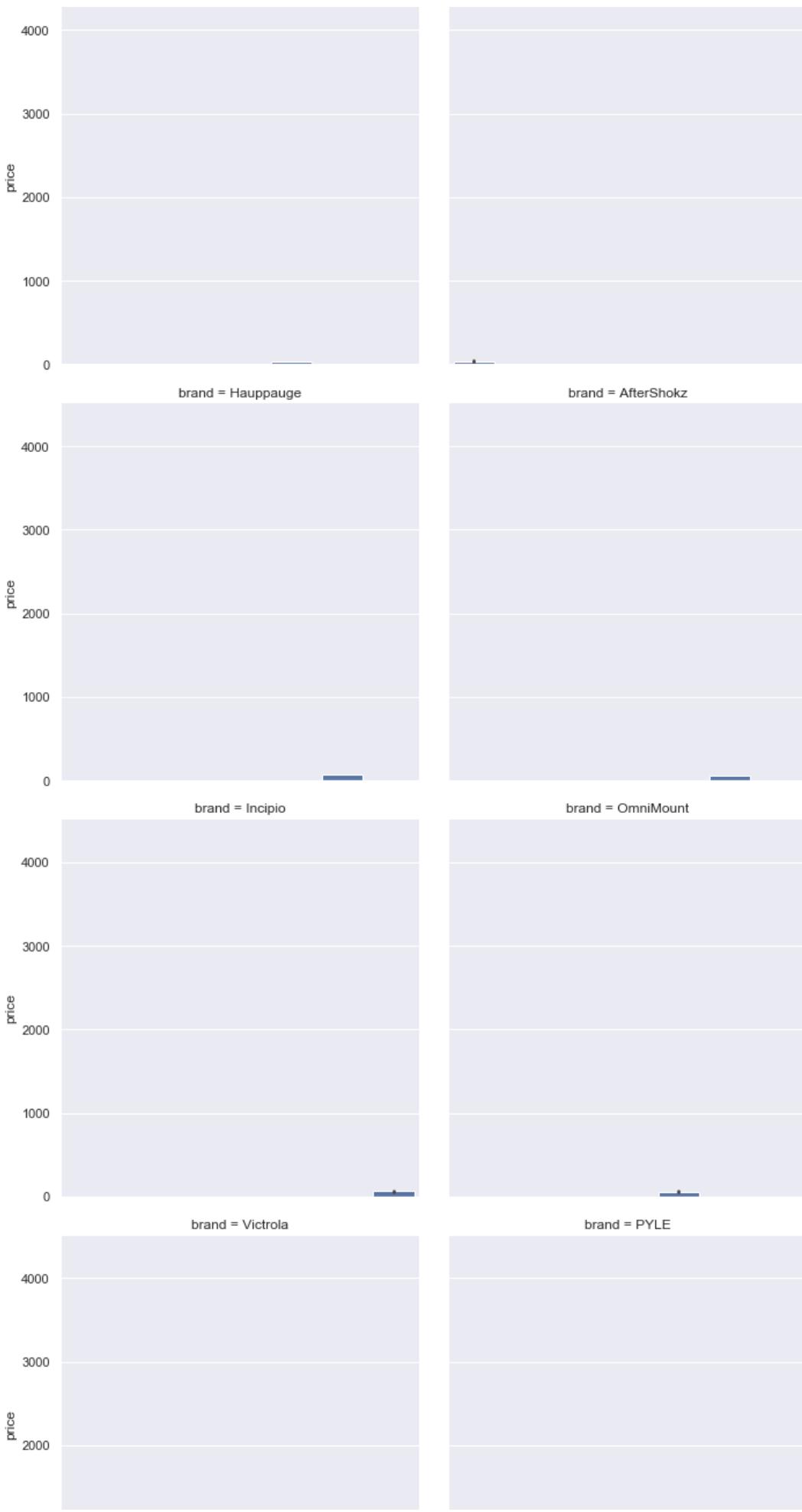


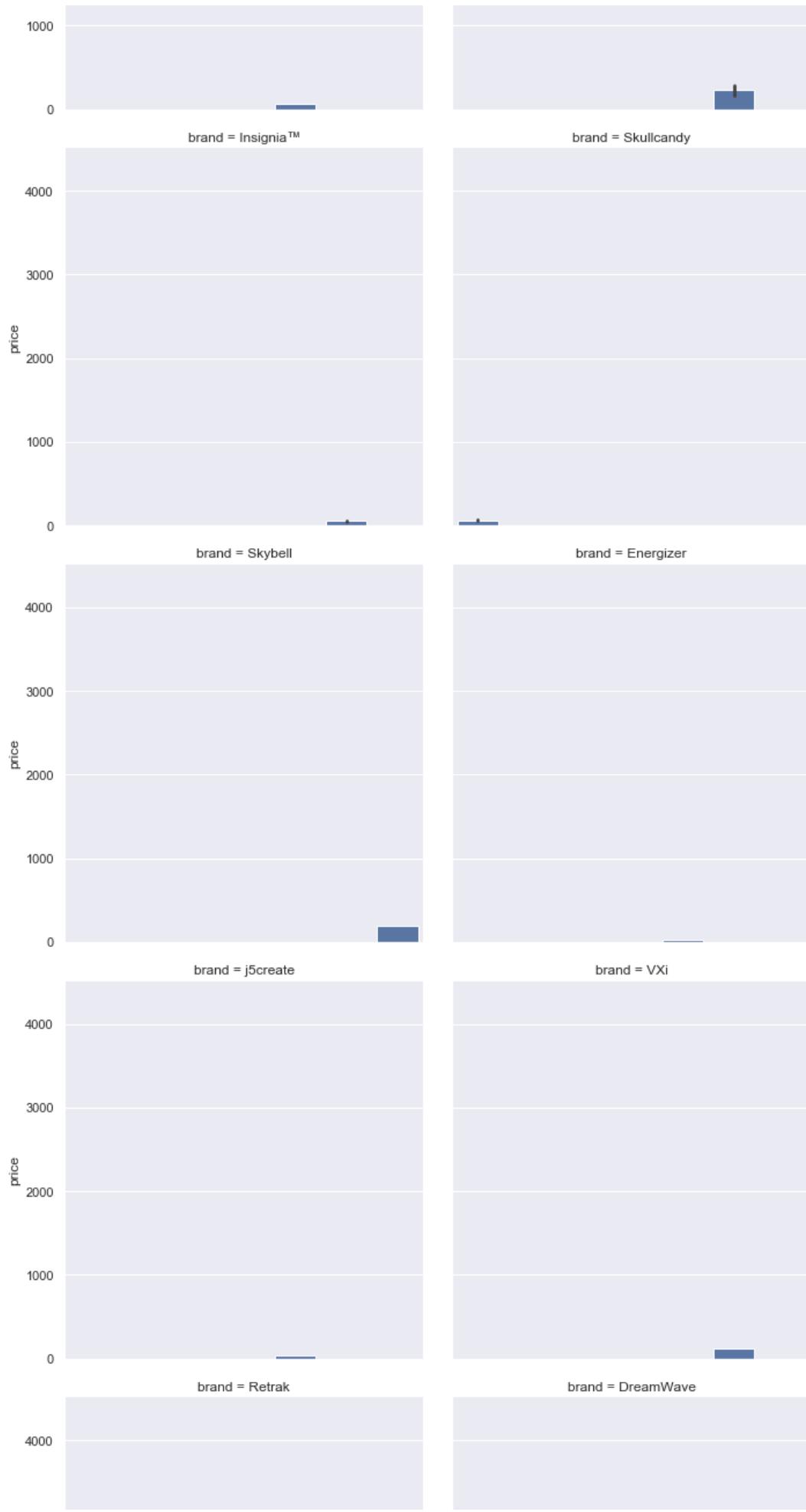


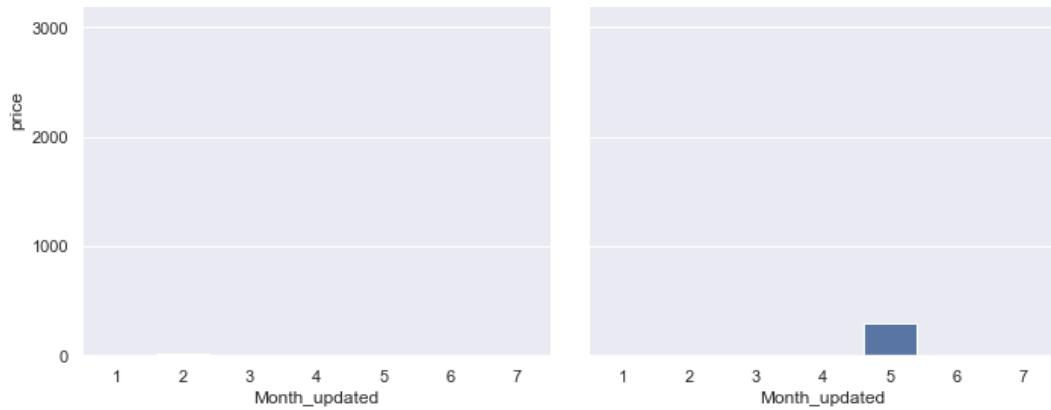












## Observation

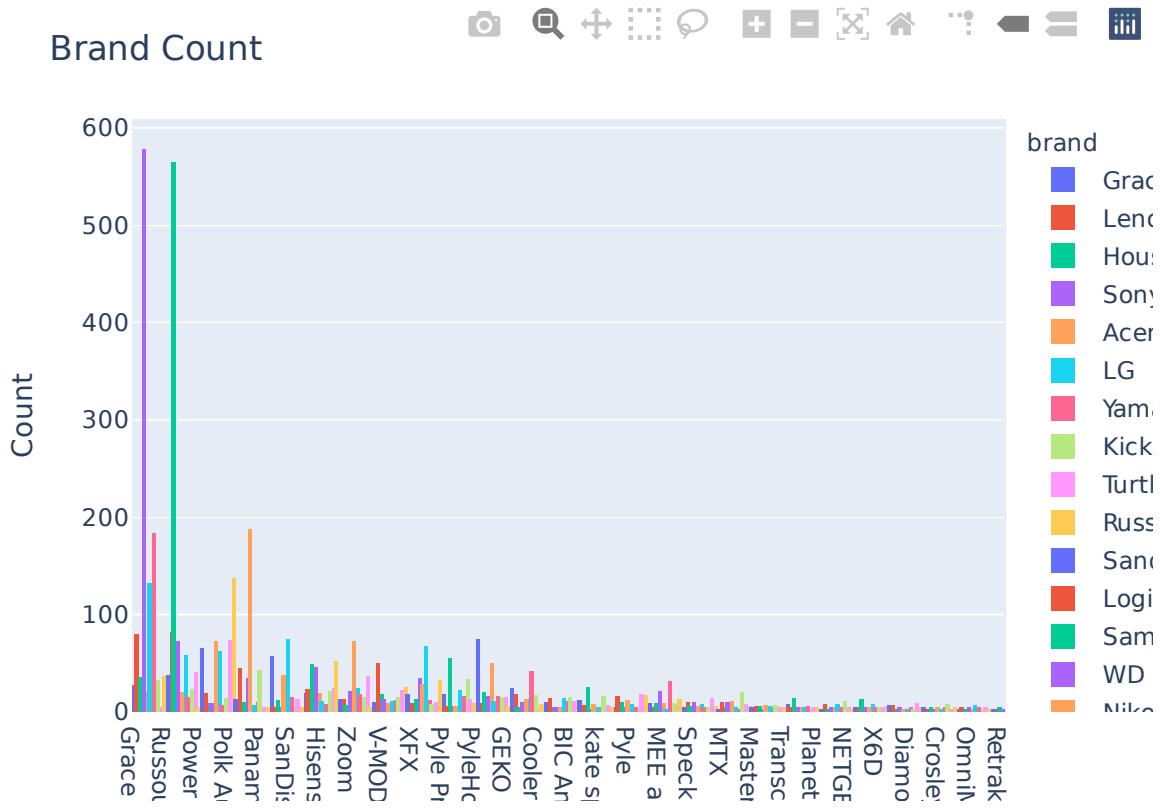
From this graph I can conclude SunBrite TV has more price than other brand  
I can also conclude that in 5th and 6th month of 2018 price is highest

For primaryCategories I will apply one-hot encoding

```
In [68]: import plotly.express as px
import plotly.graph_objects as go
```

```
In [69]: 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 [70]: histogram(dt_electronics,"brand","brand", 'Brand Count', 'Brand', 'Count')
```



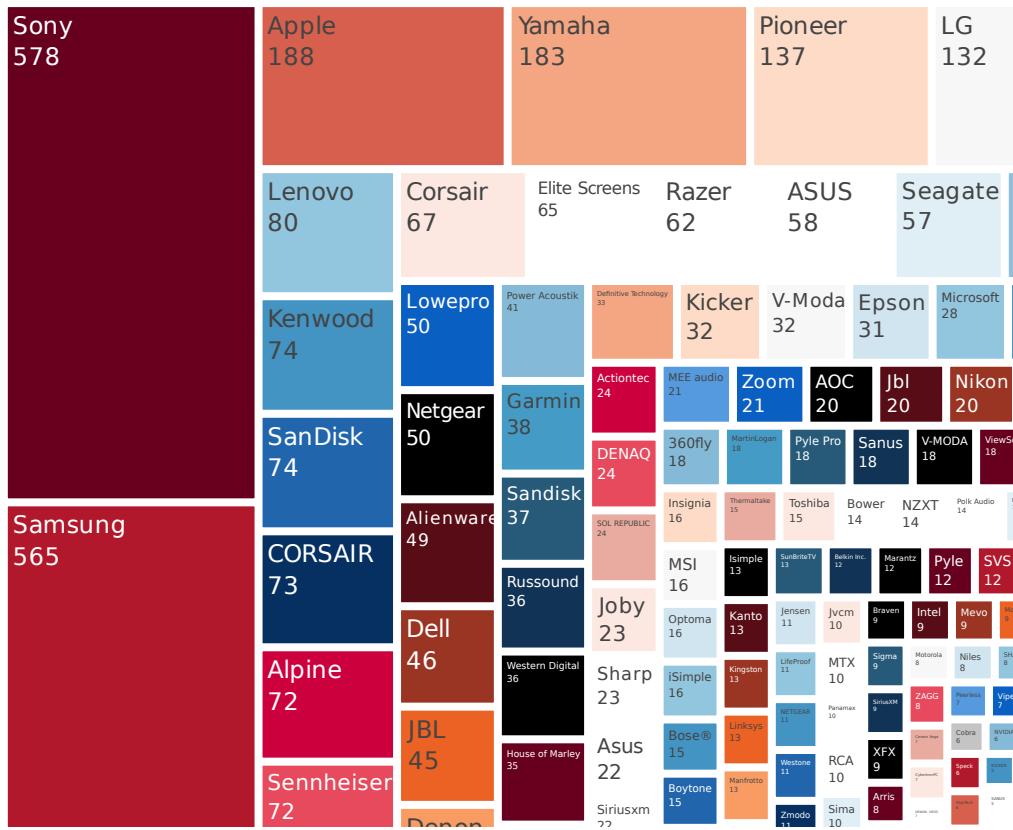
```
In [71]: Number_brand = dt_electronics['brand'].value_counts()
values = Number_brand.values
categories = pd.DataFrame(data=Number_brand.index, columns=['brand'])
categories['values'] = values
```

```
In [72]: 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 [73]: treemap(categories,'Major Sales of Brand',['brand'],categories['values'])
```



## Major Sales of Brand



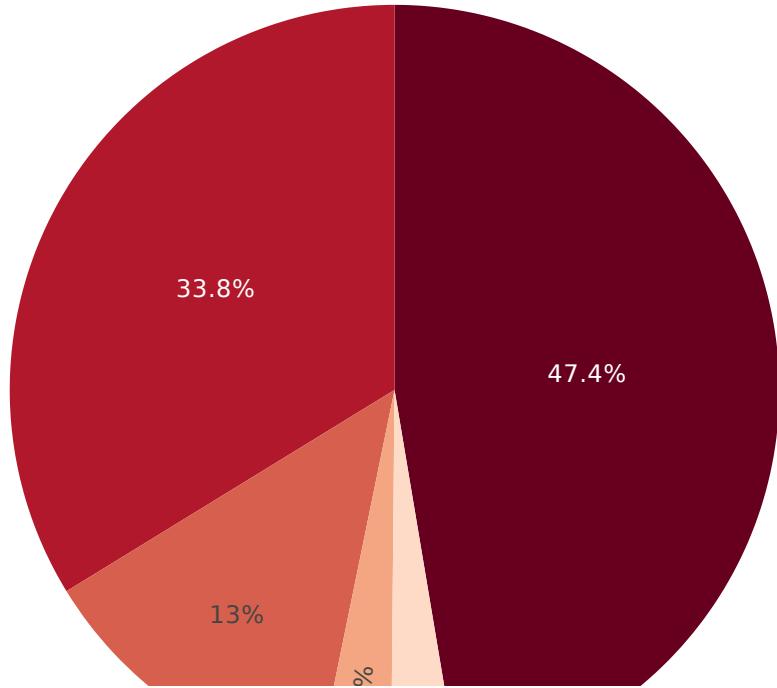
```
In [74]: def bar(categories,x,y,color,title,xlab,ylab):
    fig = px.bar(categories, x=x, y=y,
                 color=color,
                 height=400)
    fig.update_layout(
        title_text=title,
        xaxis_title_text=xlab,
        yaxis_title_text=ylab,
        bargap=0.2,
        bargroupgap=0.1
    )
    fig.show()
```

```
In [75]: bar(categories, categories['brand'][0:10],categories['values'][0:10]
           ,categories['brand'][0:10],'Top 10 brand','brand','Count')
```



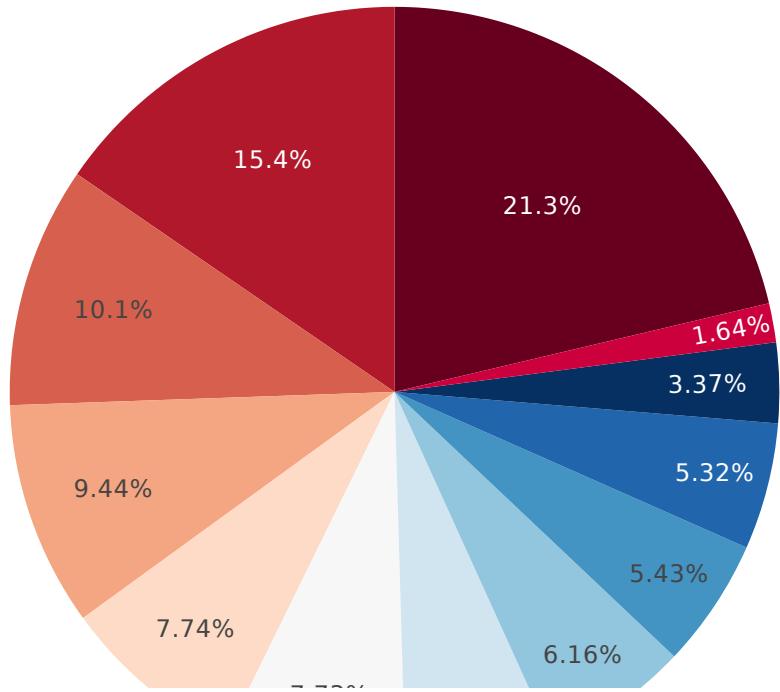
```
In [76]: Number_yeardtadded = dt_electronics['Year_dtadd'].value_counts()
years = pd.DataFrame(data=Number_yeardtadded.index, columns=["YEAR"])
years['values'] = Number_yeardtadded.values
```

```
In [77]: fig = px.pie(years, values='values', names='YEAR', color_discrete_sequence=px.colors.sequential.RdBu)
fig.show()
```



```
In [78]: Price_month = dt_electronics['Month_dtadd'].value_counts()
months = pd.DataFrame(data=Price_month.index, columns=[ "MONTH"])
months['values'] = Price_month.values
```

```
In [79]: fig = px.pie(months, values='values', names='MONTH', color_discrete_sequence=px.colors.sequential.RdBu)
fig.show()
```

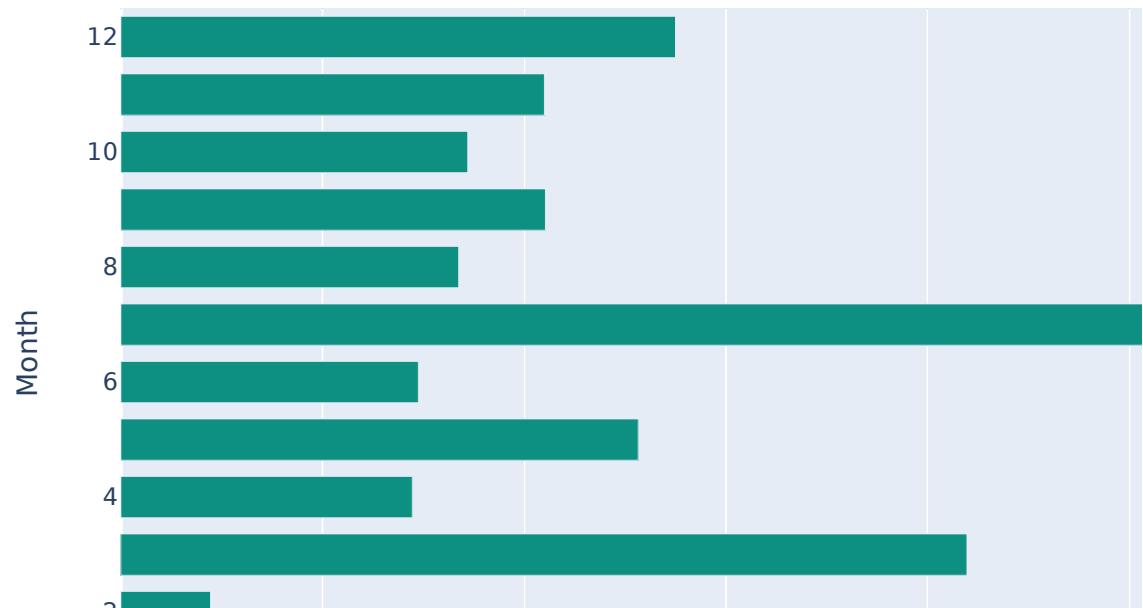


```
In [80]: fig = go.Figure(go.Bar(
    x=months['values'],
    y=months['MONTH'],
    marker=dict(
        color='rgb(13,143,129)',

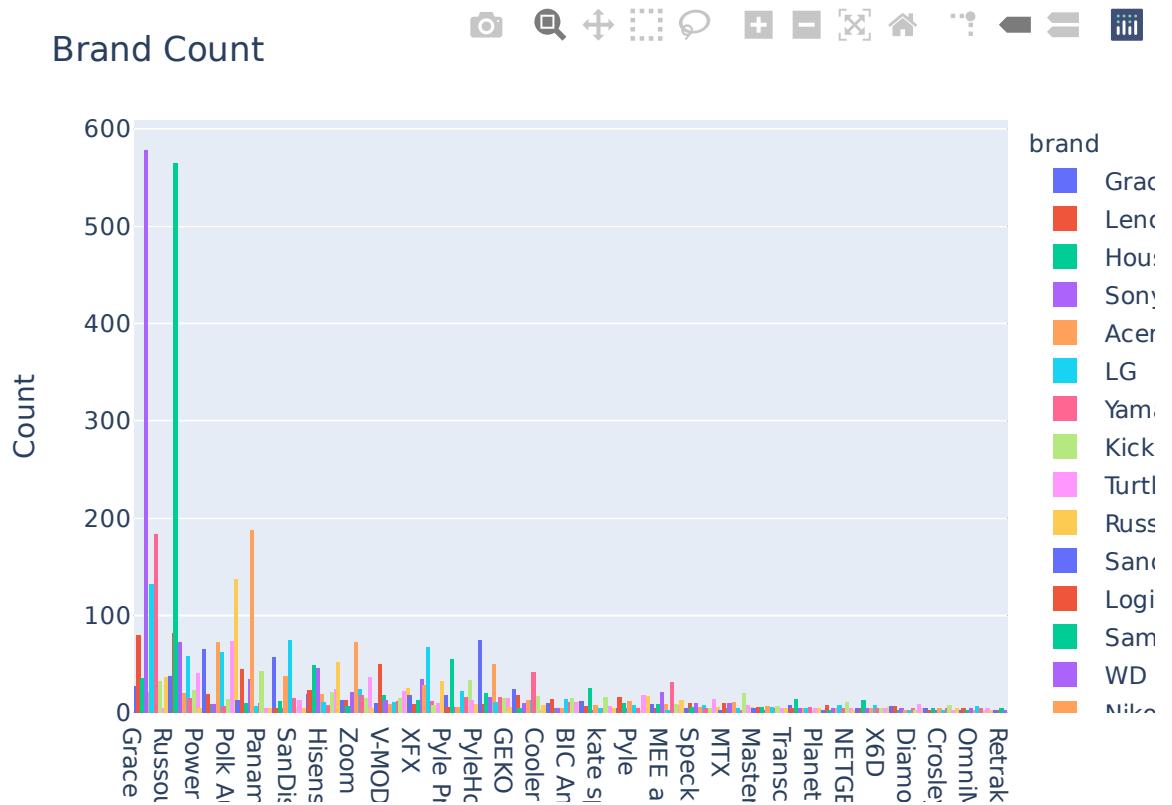
    ),
    orientation='h'))
fig.update_layout(
    title_text='Price count as per month',
    xaxis_title_text='Count',
    yaxis_title_text='Month',
    bargap=0.2,
    bargroupgap=0.1
)
fig.show()
```



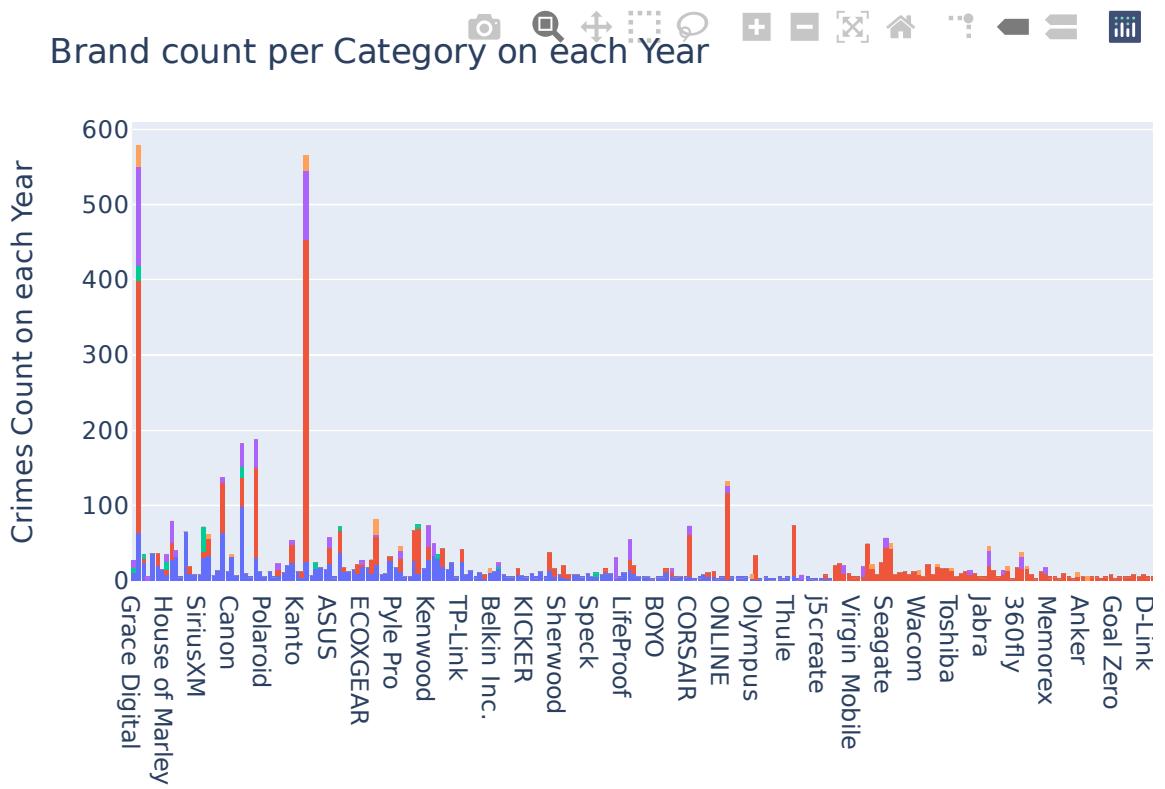
Price count as per month



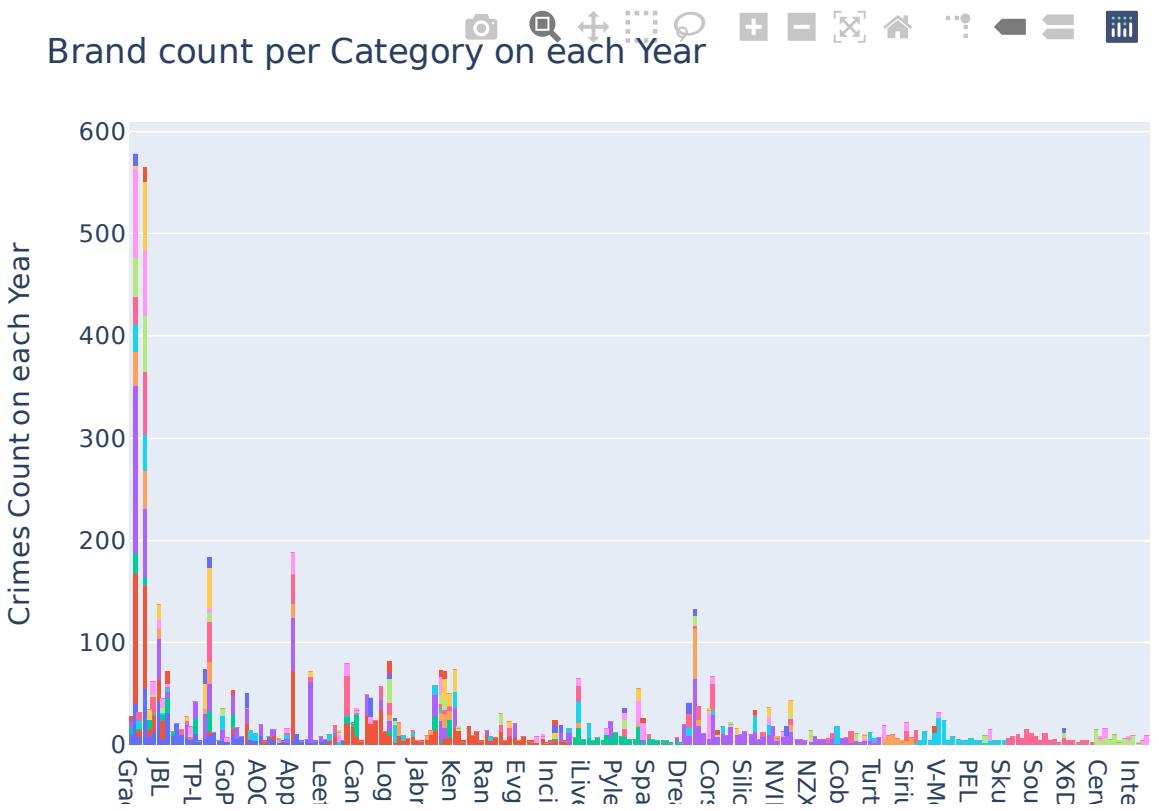
```
In [81]: histogram(dt_electronics,"brand","brand",'Brand Count','Brand','Count')
```



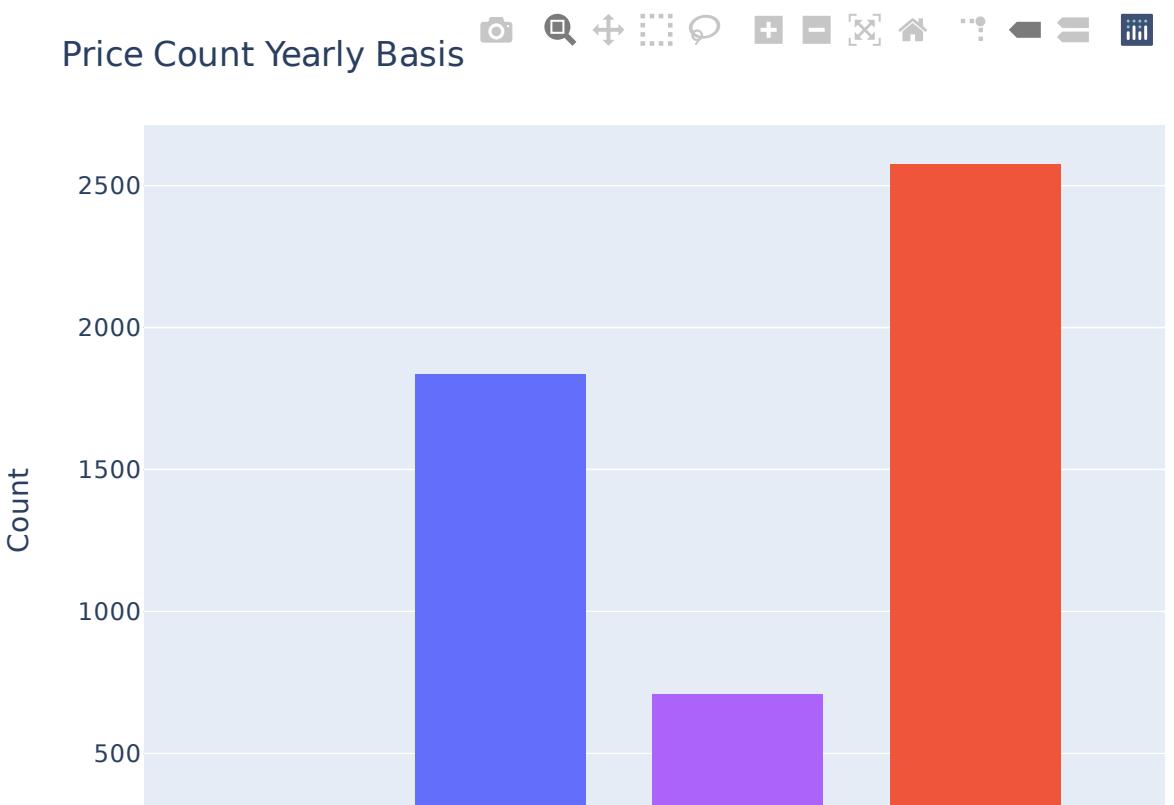
```
In [82]: histogram(dt_electronics,"brand","Year_dtadd", 'Brand count per Category on each Year','Category','Crimes Count on each Year')
```



```
In [83]: histogram(dt_electronics,"brand","Month_dtadd",'Brand count per Category  
on each Year','Category','Crimes Count on each Year')
```



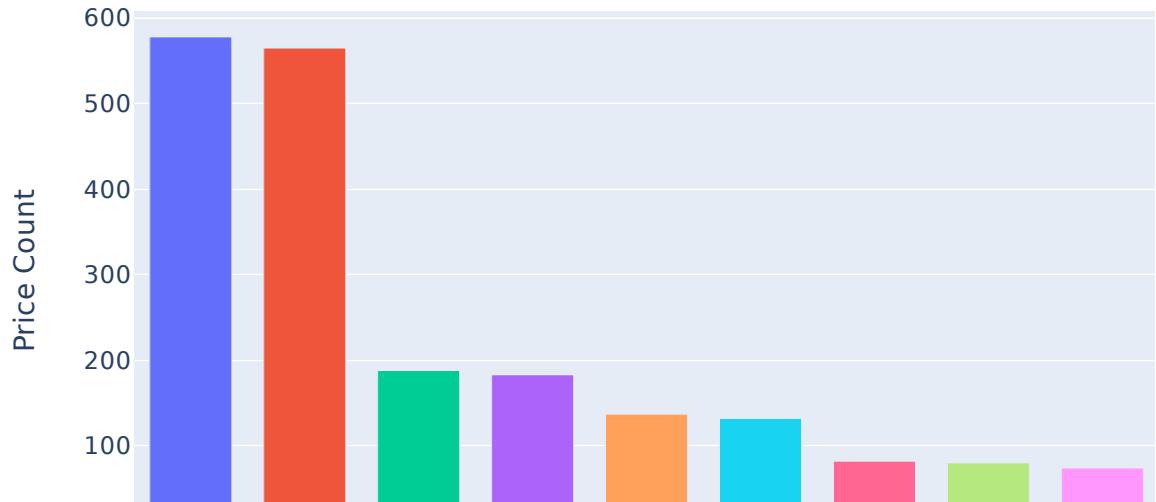
```
In [84]: histogram(dt_electronics,"Year_dtadd","Year_dtadd",'Price Count Yearly Basis','Year','Count')
```



```
In [85]: high_brand = dt_electronics['brand'].value_counts()  
brand = pd.DataFrame(data=high_brand.index, columns=[ "Brand" ])  
brand[ 'values' ] = high_brand.values
```

```
In [86]: bar(brand,brand['Brand'][0:10],brand['values'][0:10]
      ,brand['Brand'][0:10],'Top 10 brands','brands',' Price Count')
```

Top 10 brands

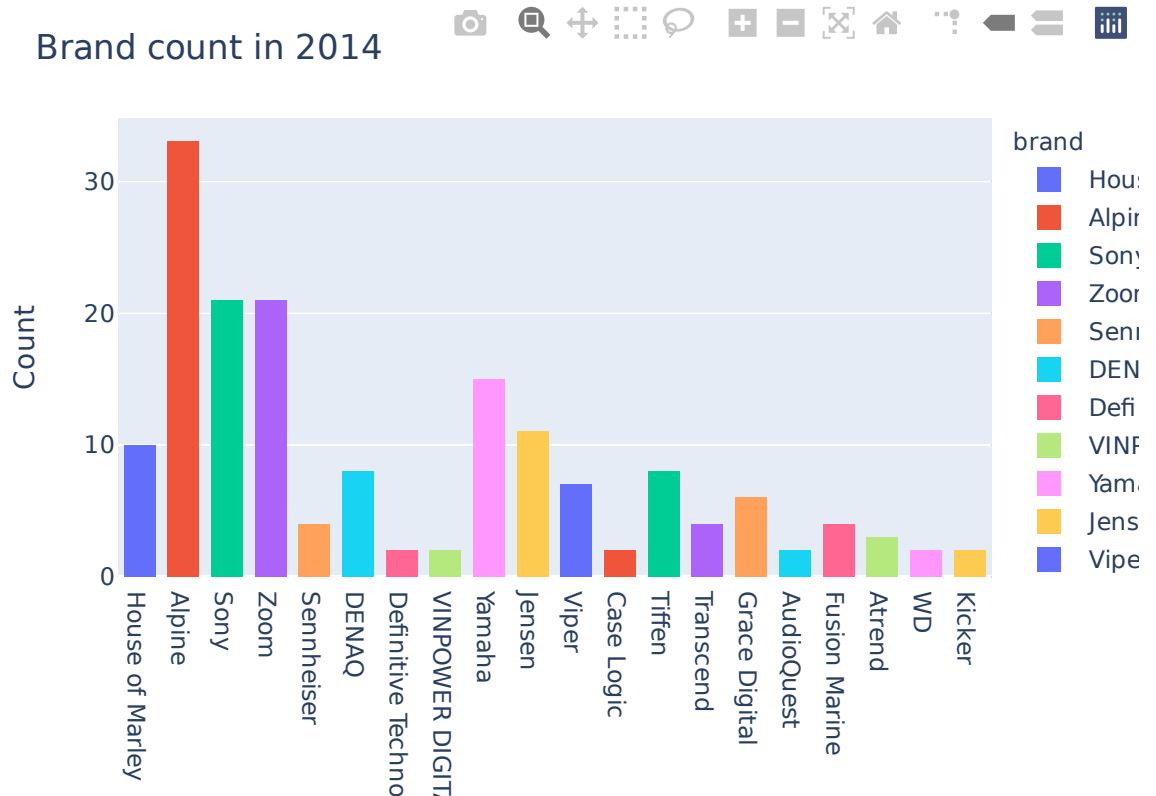


```
In [87]: dt_electronics.Year_dtadd.unique()
```

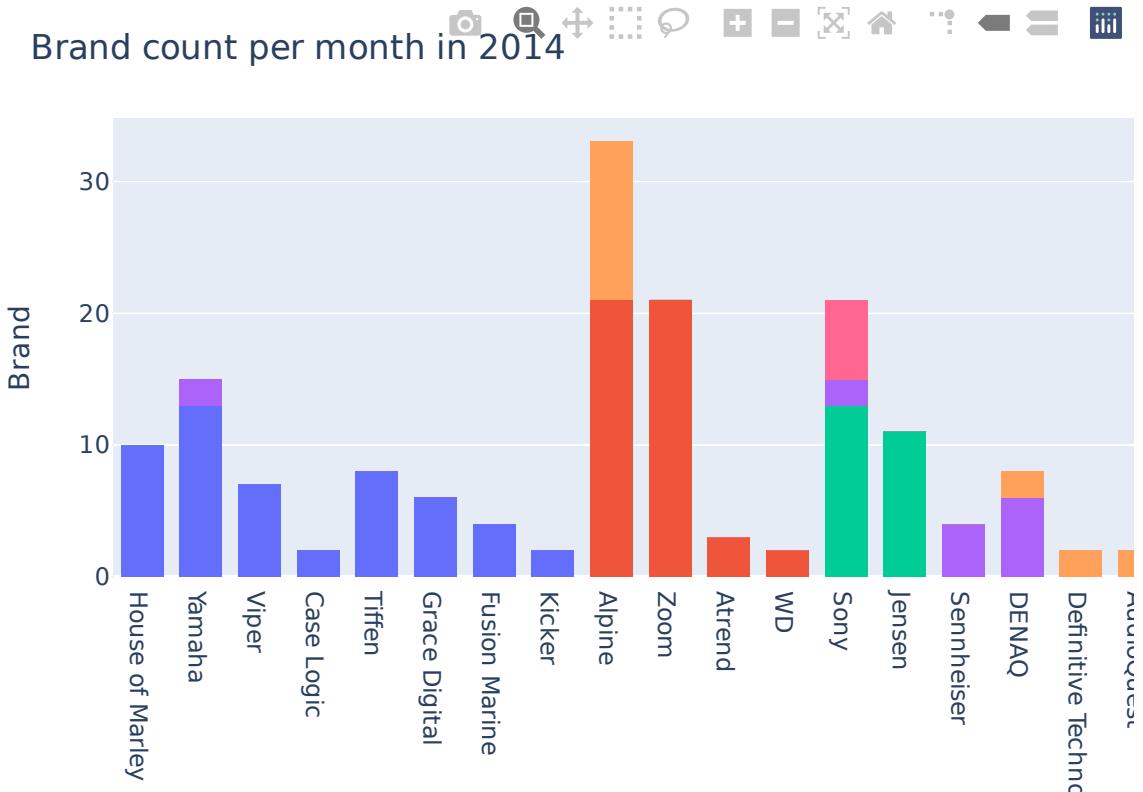
```
Out[87]: array([2015, 2017, 2014, 2016, 2018], dtype=int64)
```

```
In [88]: dt_2014=dt_electronics[dt_electronics['Year_dtadd']==2014]
```

```
In [89]: histogram(dt_2014,"brand","brand",'Brand count in 2014','Brand','Count')
```



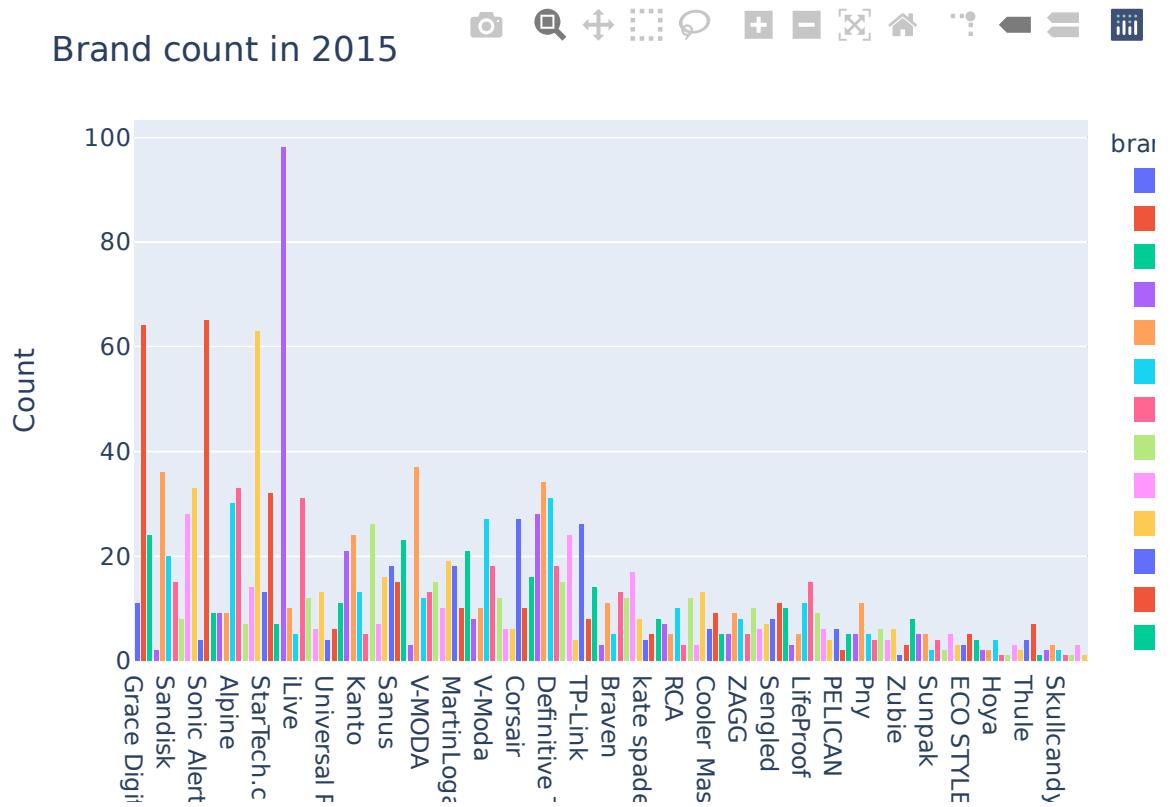
```
In [90]: histogram(dt_2014,"brand","Month_dtadd",'Brand count per month in 2014','Category','Brand')
```



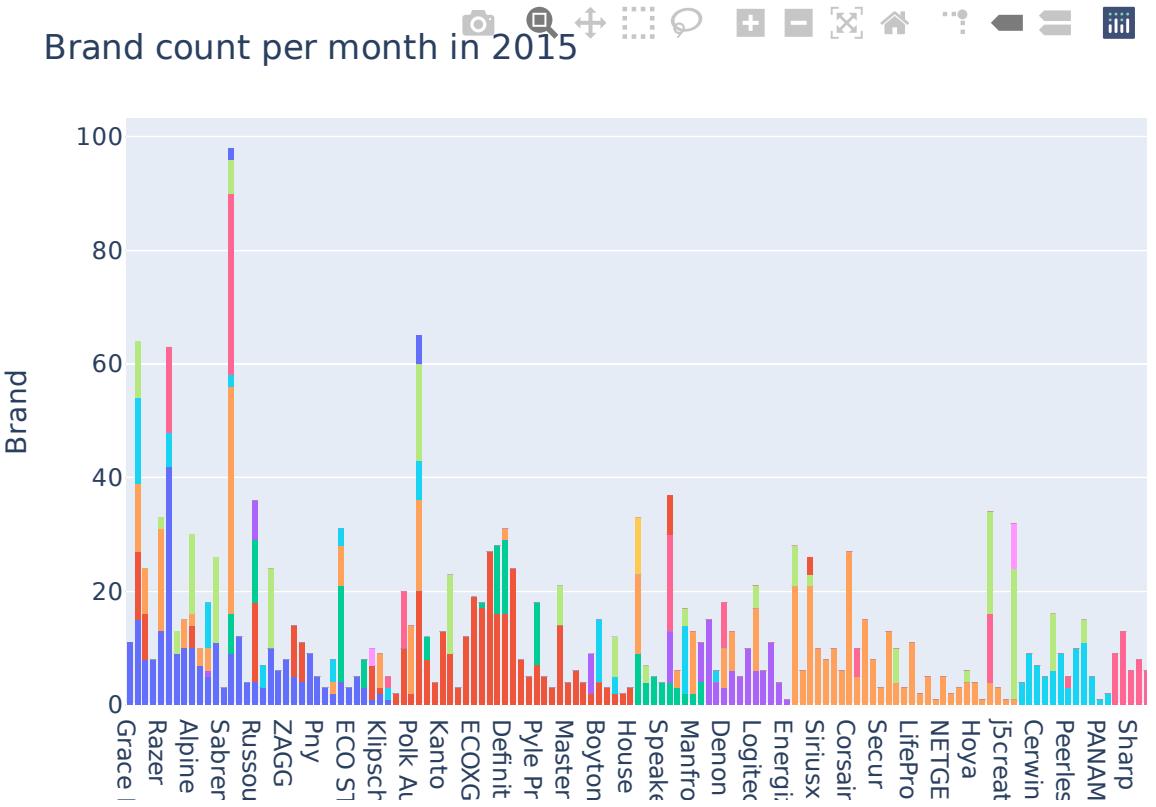
```
In Year 2014 Alpine band has the highest count
```

```
In [91]: dt_2015=dt_electronics[dt_electronics['Year_dtadd']==2015]
```

```
In [92]: histogram(dt_2015,"brand","brand",'Brand count in 2015','Brand','Count')
```



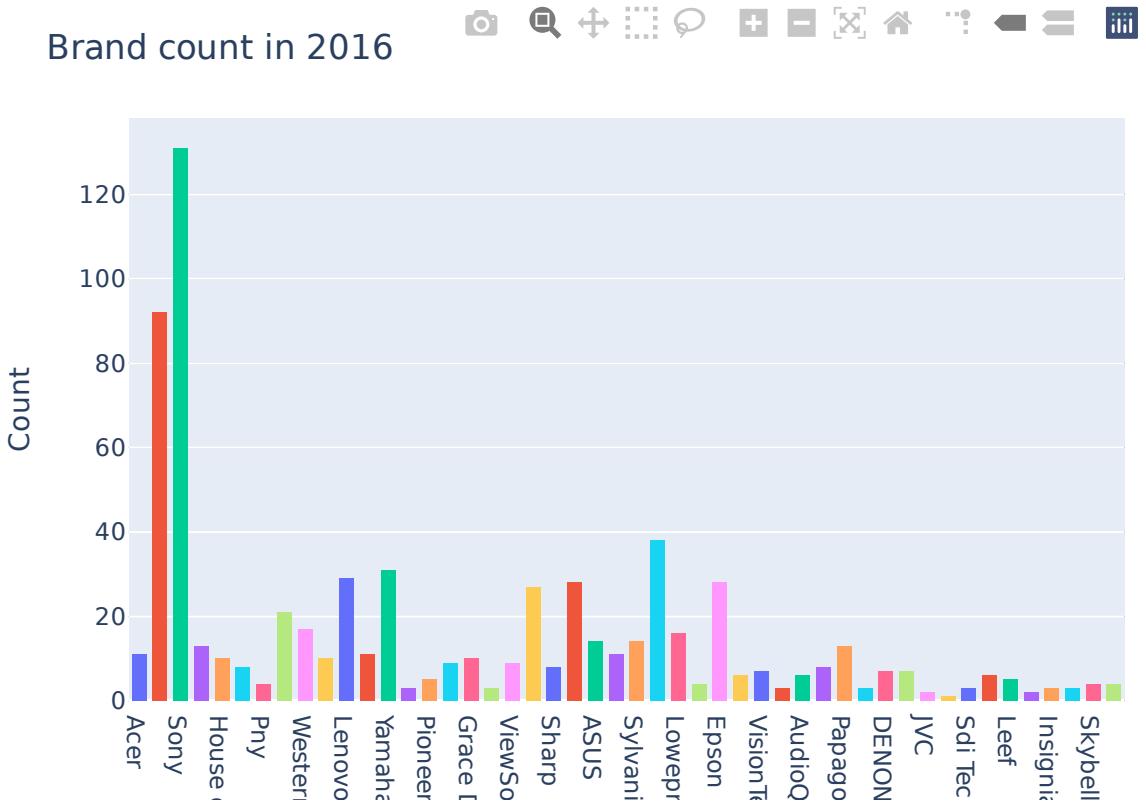
```
In [93]: histogram(dt_2015,"brand","Month_dtadd",'Brand count per month in 2015','Category','Brand')
```



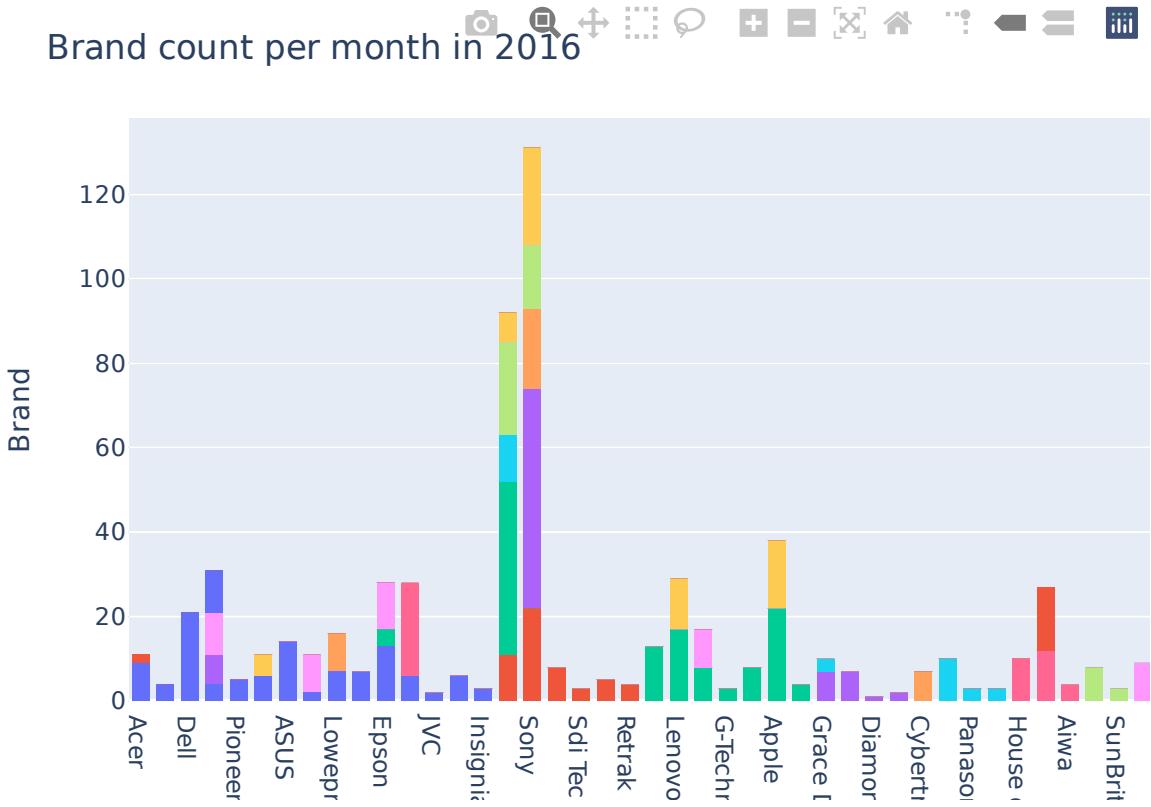
In Year 2015 Yamaha band has the highest count

```
In [94]: dt_2016=dt_electronics[dt_electronics['Year_dtadd']==2016]
```

```
In [95]: histogram(dt_2016,"brand","brand",'Brand count in 2016','Brand','Count')
```



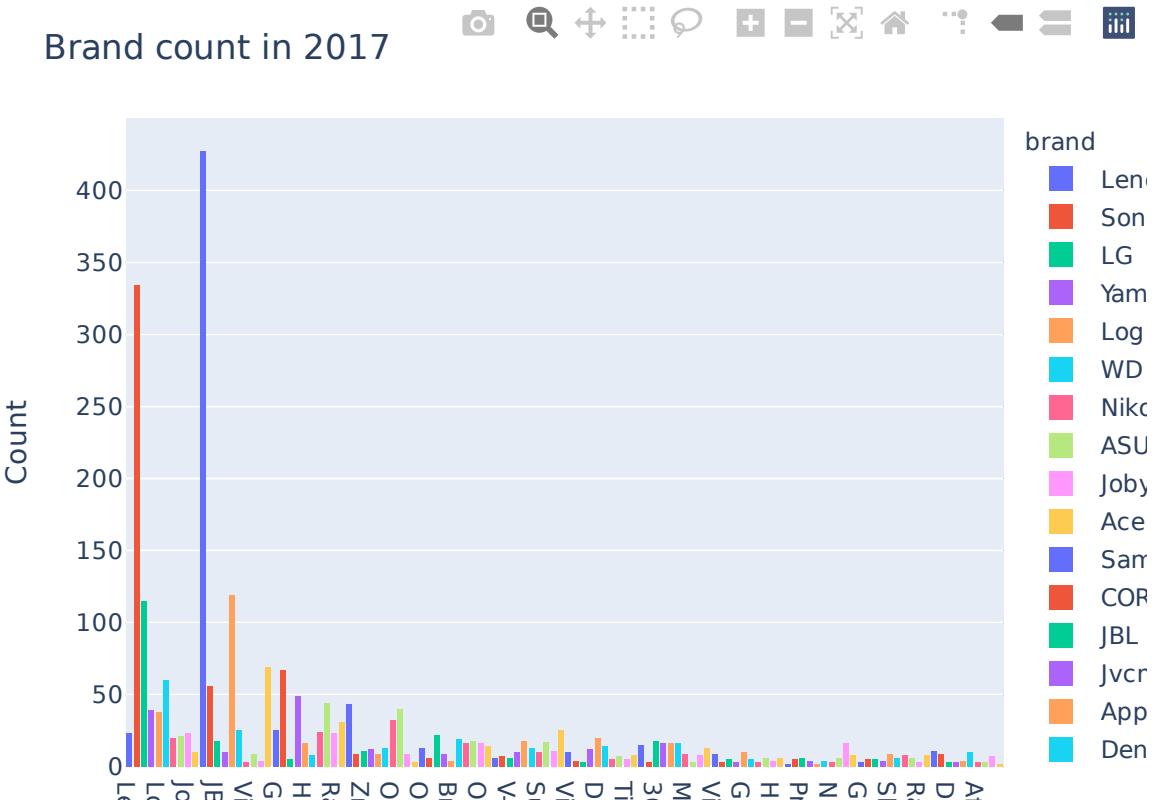
```
In [96]: histogram(dt_2016,"brand","Month_dtadd",'Brand count per month in 2016','Category','Brand')
```



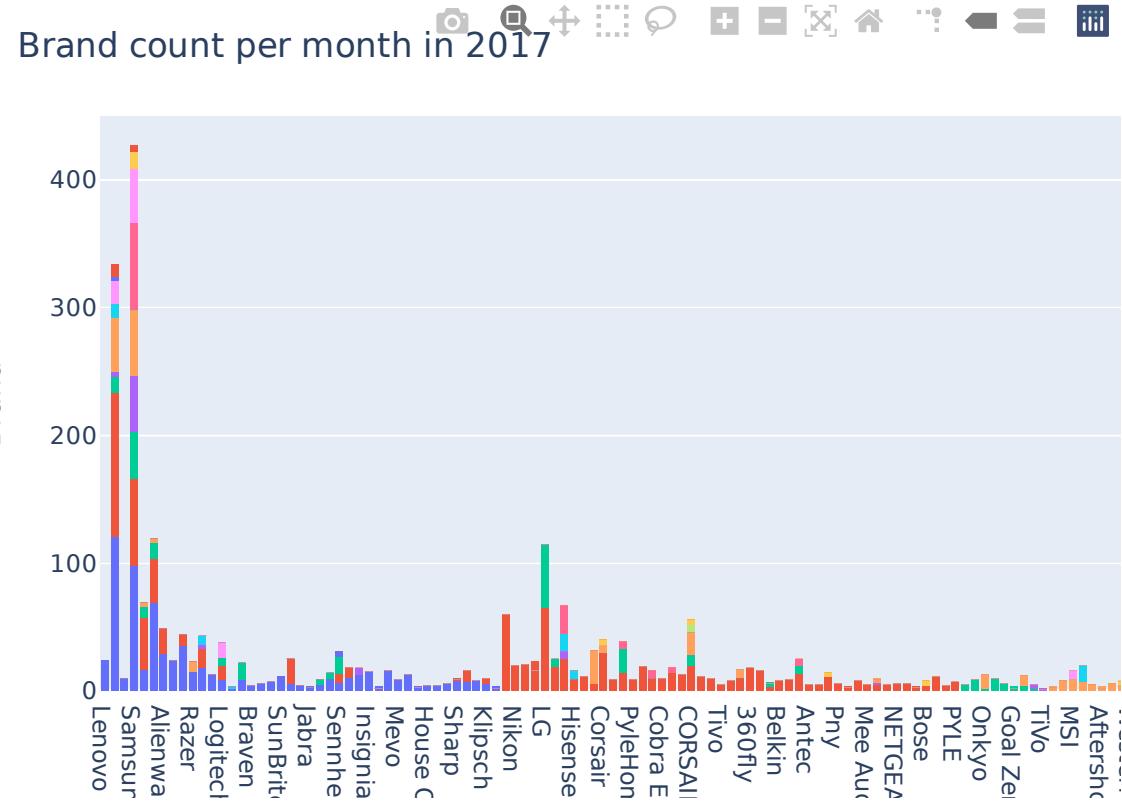
```
In Year 2016 Sony band has the highest count
```

```
In [97]: dt_2017=dt_electronics[dt_electronics['Year_dtadd']==2017]
```

```
In [98]: histogram(dt_2017,"brand","brand",'Brand count in 2017','Brand','Count')
```



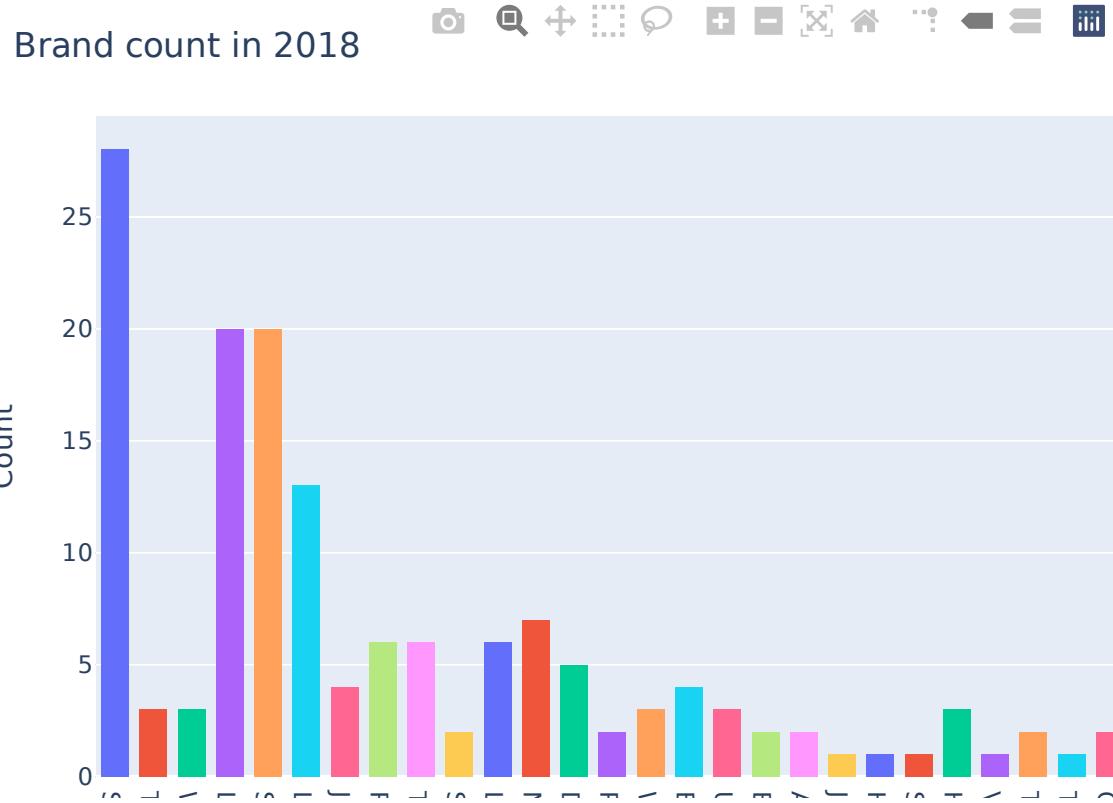
```
In [99]: histogram(dt_2017,"brand","Month_dtadd",'Brand count per month in 2017',  
'Category','Brand')
```



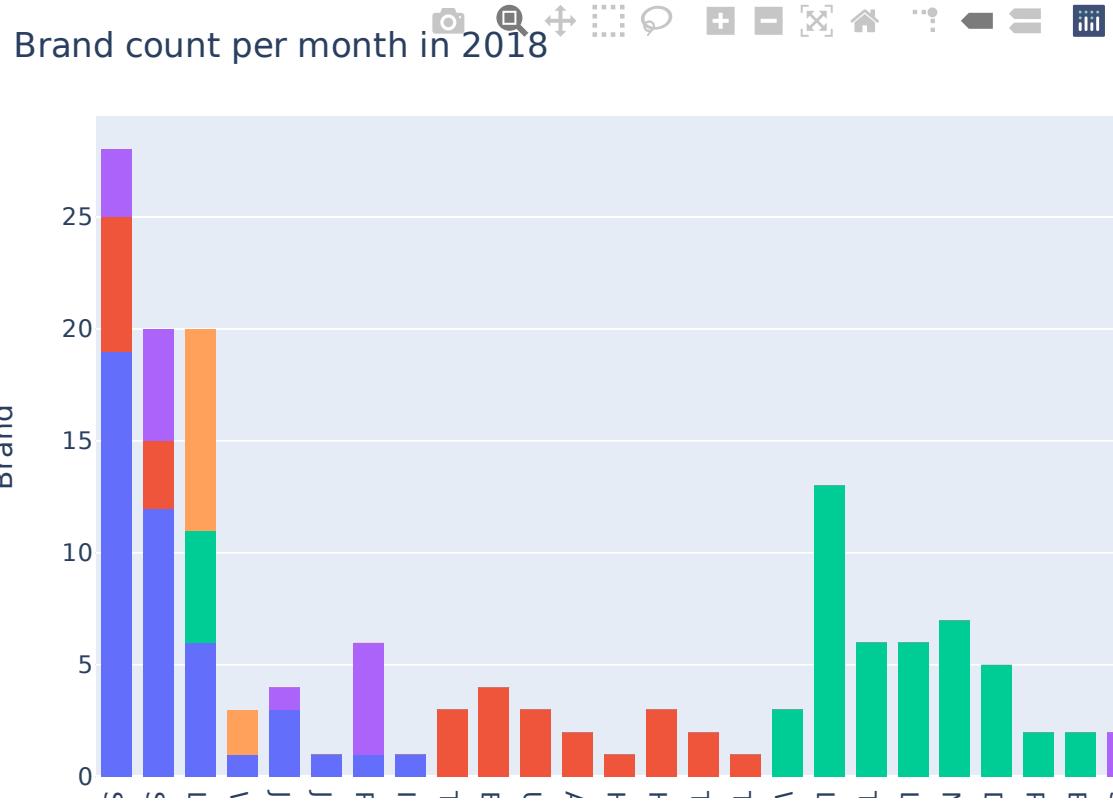
```
In Year 2017 Samsung band has the highest count
```

```
In [100]: dt_2018=dt_electronics[dt_electronics['Year_dtadd']==2018]
```

```
In [101]: histogram(dt_2018,"brand","brand",'Brand count in 2018','Brand','Count')
```



```
In [102]: histogram(dt_2018,"brand","Month_dtadd",'Brand count per month in 2018','Category','Brand')
```



In Year 2018 Sony band has the highest count

```
In [103]: dt_electronics.columns
```

```
Out[103]: 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_update
d',
       'Year_updated', 'manufacturer_count'],
      dtype='object')
```

```
In [104]: primaryCategories=pd.get_dummies(dt_electronics.primaryCategories,drop_fi  
rst=True,prefix='primaryCategories')  
dt_electronics=pd.concat([dt_electronics,primaryCategories],axis=1)  
dt_electronics.drop('primaryCategories',axis=1,inplace=True)  
dt_electronics.head()
```

Out[104]:

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

```
In [105]: dt_electronics.columns
```

```
Out[105]: Index(['pricesavailability', 'pricescondition', 'pricesdateSeen',  
'pricesmerchant', 'pricesshipping', 'pricessourceURLs', 'asins',  
'brand', 'categories', 'ean', 'imageURLs', 'keys', 'manufacturer',  
'manufacturerNumber', 'name', 'sourceURLs', 'upc', 'weight', 'pric  
e',  
'Currency_USD', 'pricesisSale_True', 'brand_ordinal_labels',  
'Month_dtadd', 'Year_dtadd', 'Month_updated', 'Year_updated',  
'manufacturer_count', 'primaryCategories_ Intel Celeron',  
'primaryCategories_ Siri Eyes Free', 'primaryCategories_Electronic  
s',  
'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 [106]: dt_electronics.drop(['categories','brand','ean', 'imageURLs', 'keys', 'ma  
nufacturer','manufacturerNumber', 'name', 'sourceURLs', 'upc', 'weight'],  
axis=1,inplace=True)
```

```
In [107]: dt_electronics.drop(['manufacturer_count'],axis=1,inplace=True)
```

## Feature Selection

**I will apply different techniques**

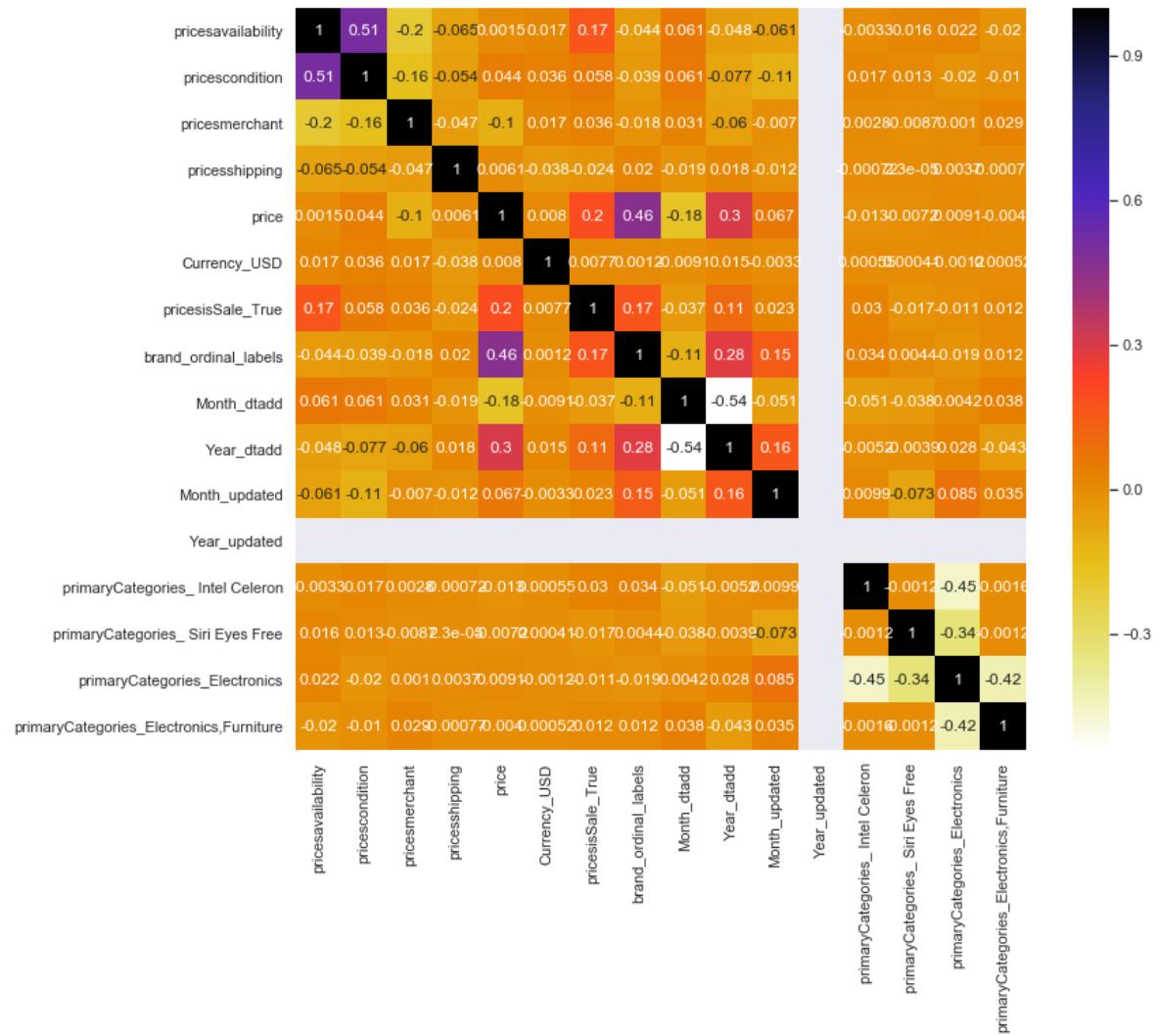
### 1. Correlation

In [108]: `dt_electronics.corr()`

Out[108]:

	pricesavailability	pricescondition	pricesmerchant	pricesshipping
<b>pricesavailability</b>	1.000000	0.508854	-0.199715	-0.065426
<b>pricescondition</b>	0.508854	1.000000	-0.156314	-0.054341
<b>pricesmerchant</b>	-0.199715	-0.156314	1.000000	-0.043529
<b>pricesshipping</b>	-0.065426	-0.054341	-0.043529	1.000000
<b>price</b>	0.001523	0.044039	0.001523	0.001523
<b>Currency_USD</b>	0.017189	0.035772	0.017189	0.017189
<b>pricesisSale_True</b>	0.165858	0.057858	0.165858	0.165858
<b>brand_ordinal_labels</b>	-0.043529	-0.038755	-0.043529	-0.043529
<b>Month_dtadd</b>	0.061236	0.061123	0.061236	0.061236
<b>Year_dtadd</b>	-0.048377	-0.077155	-0.048377	-0.048377
<b>Month_updated</b>	-0.061485	-0.113257	-0.061485	-0.061485
<b>Year_updated</b>	NaN	NaN	NaN	NaN
<b>primaryCategories_ Intel Celeron</b>	-0.003307	0.016780	0.016289	0.016289
<b>primaryCategories_ Siri Eyes Free</b>	0.016289	0.012502	0.016289	0.016289
<b>primaryCategories_Electronics</b>	0.021845	-0.019811	0.021845	0.021845
<b>primaryCategories_Electronics,Furniture</b>	-0.020494	-0.010420	-0.020494	-0.020494

```
In [109]: plt.figure(figsize=(12,10))
cor = dt_electronics.corr()
sns.heatmap(cor, annot=True, cmap=plt.cm.CMRmap_r)
plt.show()
```



```
In [110]: 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
                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 [111]: corr_features = correlation_feature(dt_electronics, 0.5)
len(set(corr_features))
```

```
Out[111]: 2
```

```
In [112]: corr_features
```

```
Out[112]: {'Year_dtadd', 'pricescondition'}
```

## 2. Finding zero variance columns

```
In [113]: zero_var=dt_electronics.var()[dt_electronics.var()==0].index.values
```

```
In [114]: zero_var
```

```
Out[114]: 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 [115]: dt_electronics.columns
```

```
Out[115]: Index(['pricesavailability', 'pricescondition', 'pricesdateSeen',
       'pricesmerchant', 'pricesshipping', 'pricessourceURLs', 'asins',
       'price', 'Currency_USD', 'pricesisSale_True', 'brand_ordinal_labels',
       'Month_dtadd', 'Year_dtadd', 'Month_updated', 'Year_updated',
       'primaryCategories_ Intel Celeron', 'primaryCategories_ Siri Eyes
       Free',
       'primaryCategories_Electronics',
       'primaryCategories_Electronics,Furniture'],
      dtype='object')
```

```
In [116]: dt_electronics.drop('pricesdateSeen',axis=1,inplace=True)
```

```
In [117]: dt_electronics.drop(['pricessourceURLs','asins'],axis=1,inplace=True)
```

```
In [118]: dt_electronics.drop(['Year_dtadd','pricescondition','Year_updated'],axis=
       1,inplace=True)
```

# Model Development

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

## Performance Metric

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

```
In [119]: x=dt_electronics.drop('price',axis=1)  
y=dt_electronics['price']
```

Splitting the dataset into train and test dataset

```
In [120]: x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3,random_state=42)
```

```
In [121]: x_train.shape,y_train.shape
```

```
Out[121]: ((3805, 12), (3805,))
```

## Linear Regression

```
In [122]: 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 [123]: print('R2 square for train',r2_score(y_train,model_pred_tr))  
print('R2 square for test',r2_score(y_test,model_pred_test))
```

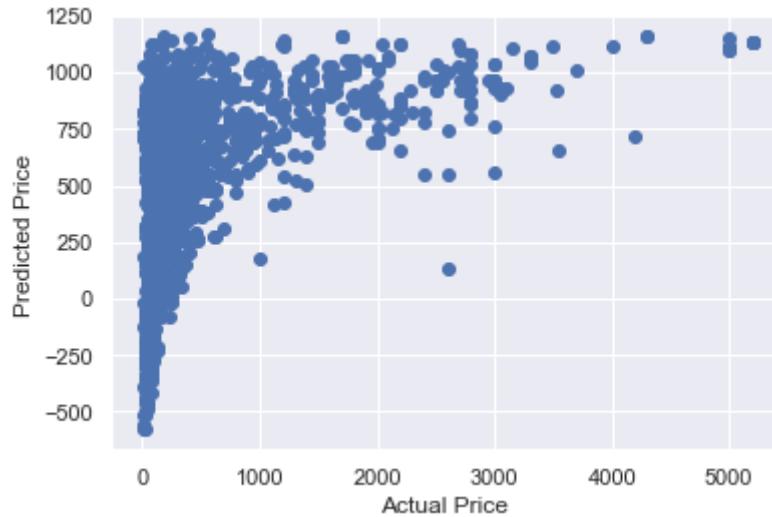
```
R2 square for train 0.25029309674174804  
R2 square for test 0.25411079403472836
```

## 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 [124]: plt.scatter(y_test,model_pred_test)  
plt.xlabel('Actual Price ')  
plt.ylabel('Predicted Price ')
```

```
Out[124]: Text(0, 0.5, 'Predicted Price ')
```



## Random Forest

```
In [125]: 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\fores  
t.py:245: FutureWarning:
```

The default value of n\_estimators will change from 10 in version 0.20 to 100 in 0.22.

```
In [126]: print('R2 square for train',r2_score(y_train,model_rand_train))
print('R2 square for test',r2_score(y_test,model_rand_test))
```

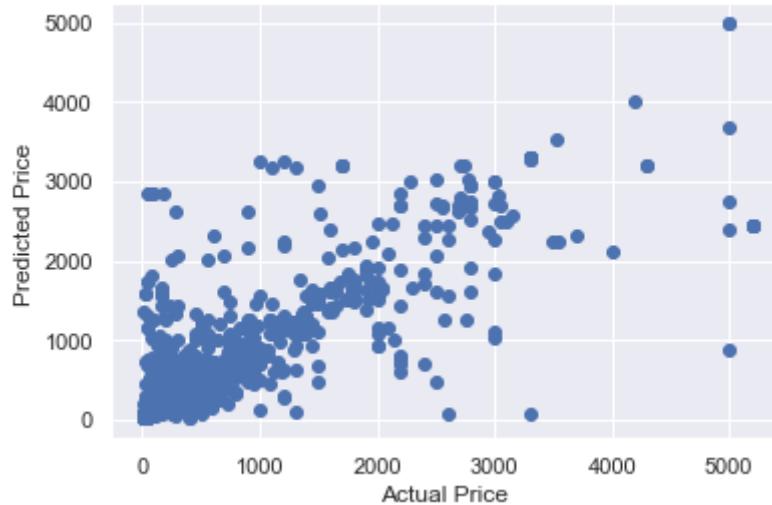
```
R2 square for train 0.8590821598025388
R2 square for test 0.6690774022260699
```

## 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 [127]: plt.scatter(y_test,model_rand_test)
plt.xlabel('Actual Price ')
plt.ylabel('Predicted Price ')
```

```
Out[127]: Text(0, 0.5, 'Predicted Price ')
```



## Decision Tree

```
In [128]: 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 [129]: print('R2 square for train',r2_score(y_train,dct_mod_train))
print('R2 square for test',r2_score(y_test,dct_mod_test))
```

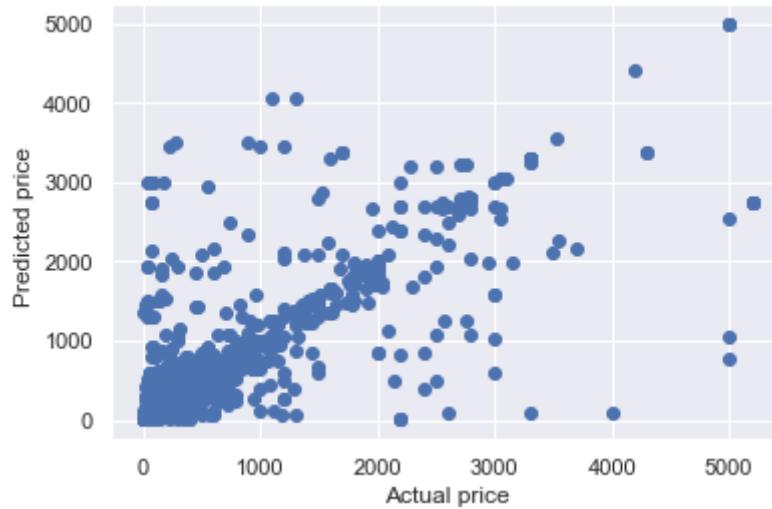
```
R2 square for train 0.887933143003788
R2 square for test 0.5775396171730032
```

## 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 [130]: plt.scatter(y_test,dct_mod_test)
plt.xlabel('Actual price')
plt.ylabel('Predicted price ')
```

```
Out[130]: 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 [ ]: