


The dataset includes 3,150 observations and 5 variables ('rating', 'date', 'variation', 'verified_reviews', 'feedback')

	rating	date	variation \
0	5	31-Jul-18	Charcoal Fabric
1	5	31-Jul-18	Charcoal Fabric
2	4	31-Jul-18	Walnut Finish
3	5	31-Jul-18	Charcoal Fabric
4	5	31-Jul-18	Charcoal Fabric

verified reviews feedback

Consumer Ratings ranges from 1 to 5, just like feedback variable, positive ratings (4, 5) are majority.

```
dff['rating'].hist(bins = 5)
sns.plotlib.axes_subplots.AxesSubplot at 0x266a5659c0e0
```



Rating	Frequency
1	0
2	0
3	0
4	2000
5	2500

```
array(['Charcoal Fabric ', 'Walnut Finish ', 'Heather Gray Fabric ',
       'Sandstone Fabric ', 'Oak Finish ', 'Black', 'White',
       'Black Spot', 'White Spot', 'Black Show', 'White Show',
       'Black Plus', 'White Plus', 'Configuration: Fire TV Stick',
       'Black Dot', 'White Dot'], dtype=object)
```

Country	Publications
Argentina	10
Brazil	100
Canada	100
Germany	100
India	100
Japan	100
Korea	100
Mexico	100
New Zealand	100
Norway	100
Peru	100
Russia	100
Saudi Arabia	100
Switzerland	100
USA	100
UK	100

information for sentiment analysis. So we are going to drop date variable as well. Lastly, we saw from above chart that neither feedback nor rating varies by product variation, so we are going to drop 'variation' as well.

```
df = df.drop(['date', 'rating','variation'],axis=1)
df.head()
```

	verified_reviews	feedback	length
0	Love my Echo!	1	13
1	Loved it!	1	9
2	Sometimes while playing a game, you can answer...	1	195
3	I have had a lot of fun with this device. My 4	1	172

```
from sklearn.feature_extraction.text import CountVectorizer
vectorizer = CountVectorizer()
review_vectorized = vectorizer.fit_transform(df['verified_reviews'])
```

[illegible][illegible]

ceilings, celebs, cell/cellphone, cent, center, certain, certainly, certified, chackki, chair
 'chak', challenge, 'challenged', champ, 'chance', change, 'changed', 'changer', 'changes', 'changing', 'chain'
 'l', channels, 'characteristics', 'charge', 'chargeable' check 'checking', charlotte, 'child', 'charmed',
 'chill', 'chilling', chat, 'chatting', cheap, cheater, 'cheapest', 'check', 'checked', 'checking', 'child', 'children',
 idler, 'chocolate', choice, choices, choose, choosing, choppy, 'church', 'churches', 'christmas',
 'chromobok', chromecast, 'circle', citizens, 'city', clapper, clarity, classes, 'classic', classical,
 classroom, clean, cleaner, cleaning, clear, clearer, clearly, 'click', 'clips', clock, 'clock'

[illegible][illegible][illegible][illegible][illegible]

live, 'lyric, 'lyrical, 'lyrics, 'mac, 'machine, 'machines, 'maddening, 'made, 'madlib, 'magically, 'make
ed, 'main, 'mainly, 'mainstream, 'maintain, 'maintaining, 'majel, 'majes, 'major, 'nake, 'makes, 'making
'makings, 'male, 'maison, 'mana, 'nann, 'manage, 'management, 'manly, 'manoeuvre, 'manners, 'manual, 'narrative, 'narrow, 'narrowly, 'narrowness, 'narrowness, 'narrowness, 'narrowness, 'narrowness, 'narrowness
'nave, 'natch, 'matched, 'material, 'matter, 'maximize, 'may, 'maybe, 'sb, 'se, 'seal, 'sealing, 'meaningful, 'means
'neant, 'media, 'medical, 'medications, 'medicore, 'meditation, 'medium, 'mel, 'member, 'members, 'members
ip, 'neary, 'ention, 'mentioned, 'sensu, 'nere, 'message, 'messages, 'messaging, 'messsed, 'net, 'netro

[illegible]

est, press, presumably, prettier, pretty, prevent, prevents' preview, previous, previously, pre-
ceded, priced, prices, pricey, pricing, primarily, primary, prime, primed, prideday, print, prior, private,
prize, prize-winning, probably, probable, probability, produce, production, products, product's
ucts, productsand, profiles, program, programming, programmed, programming, programs, project, pro-
jector, projects, promised, promoting, promotion, prompt, prompts, proper, properly, propo-
sals, pros, protected, protection, protocol, prove, proving, provide, provided, provider, provid-
es, provides, providing, psychological, puerile, pull, pulling, pulsate, pulsed, punched, punchy, push,
pushed, pushing, psychical, psychics, psychopaths, psychoses, pursue, pursuing, pursuit, pursued, pursues,

enplaced, replacement, recedes, replacing, replied, replying, report, reported, reports, reports
rn, reputation, request, requesting, requests, require, required, requires, research, researched
researching, resembling, resection, reset, resetting, resist, resistant, resistance, resolved
resolving, resolve, resolves, resolutions, resolve, resolved, resolves, resolutions, resolved
rest, restart, restriction, restrictive, result, results, resume, retired, return, returned
turn, returning, review, reviewing, reviews, revise, rewards, rid, rides, ridiculous, ridicu-
sly, right, ring, rings, rivers, road, rock, rocks, rock, roll, room, rooms, roomy, roomy,
turn, returning, review, reviewing, reviews, revise, rewards, rid, rides, ridiculous, ridicu-

[illegible][illegible][illegible]

```
reviews = pd.DataFrame(review_vectorized.toarray())
```

[illegible]

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.30, random_state=5)

print("X_train", X_train.shape)
print("X_test", X_test.shape)
print("y_train", y_train.shape)
```

```
RandomForestClassifier(max_features='auto',
                       min_samples_leaf=1,
                       min_samples_split=2,
                       min_weight_fraction_leaf=0.0,
                       n_estimators=100,
                       verbose=0,
                       warm_start=False)
```

5. Random Forest - Evaluating

```
print(classification_report(y_test, y_predict_test))
```