Rafael Hernandez
Adidas vs Nike: Report and Conclusions

Data and its source: https://www.kaggle.com/kaushiksuresh147/adidas-vs-nike

DATA TERMS

Product Name: Name given to product by Adidas or Nike
Product ID: Unique product id given to the product.
Listing Price: Price of the product as listed
Sale Price: Price at which the product was sold.
Discount: Discount provided on the product
Brand: Brand of the product
Description: Description of the product
Rating: Rating provided by the product by users
Reviews: Number of reviews provided on the product
Last Visited: Last visit on the product by a customer.

Description of data exploration and data cleaning steps

The data compares Adidas vs Nike shoes. To clean the data, we narrowed the focus of the data and only kept the columns containing Product Name, Brand, Rating and Reviews. The data had listing price, sale price and discount columns but it lacked actual sales numbers. The data does not tell us how many shoes were sold and a decision was made to focus on information that tells us how consumers felt about the shoes which we can analyze through the Rating and Reviews columns. We kept Product Name and Brand as identifiers and worked with Rating and Reviews columns. We kept the Rating column as float64 and kept the Reviews column as int64. The data was imported and converted into a pandas data frame. The original data frame was then split into an Adidas data frame which listed the type of shoes and the Rating and Reviews data and a Nike data frame, which did the same. The objective is to compare Adidas vs Nike within the original data frame and by breaking them into their own data frames to compare statistics.

Exploration

The first exploration involved the mean of the overall data. This information gave us a baseline of the overall data which was used to try and answer the third and final questions.

```
#mean of Rating and Reviews in original data frame
best_shoesdf.mean()
```

```
Rating      3.242105
Reviews    40.551714
dtype: float64
```

COMPARISON QUESTIONS

1.Which brand is more highly rated?
2.Which brand averages more reviews?
3.Which brand makes better shoes?

Discoveries

One of the early discoveries that was made was the variance in length of shoes between Adidas and Nike. Comparing the length of the Adidas and the Nike data frames discovered that the data set has 2624 Adidas shoes and 643 Nike shoes. Apparently, Adidas makes a lot more shoes than Nike does.

```
] #comparing length of nike and adidas dataframes
  len(adidasdf)
  len(nikedf)
  print("The length of adidas data frame is " ,len(adidasdf))
  print("The length of nike data frame is " , len(nikedf))

  The length of adidas data frame is  2624
  The length of nike data frame is  643
```

To answer question number 1: **Which brand is more highly rated?** We used pandas .mean() function. We labeled the values as adidas_rating and nike_rating and took the mean from the "Ratings" column for the adidas and nike data frame. The highest rating a shoe can get is a 5.0 and the lowest is 0.0. Adidas shoes are more highly rated than Nike shoes. The print statement below reveals that Adidas averages a higher rating than Nike.

```
The average rating for adidas shoes is  3.37
The average rating for nike shoes is  2.73
```

To answer question number 2: **which brand averages more reviews?** We use .mean() function. We gave the value of adidas_reviews and nike_reviews and took the mean from the "Reviews" column in each data frame.

```
The average number of reviews for adidas shoes are  48.71
The average number of reviews for nike shoes are  7.18
```

There is a significant gap in the average number of reviews for Adidas than for Nike. What this tells us is that there is a higher intensity to express how people feel about Adidas shoes than Nike shoes. Adidas shoe owners are more willing to express themselves and those expressions can be negative or positive, but if we combine the information, we got in the first question and second question we can begin to deduce that people believe that Adidas makes a better shoe.

The third question asks **which brand makes better shoes?** As stated above we can begin to deduce that Adidas makes better shoes from the average rating and the average number of reviews. We wanted a list of the shoes who ranked in the 94th-95th percentile (the top 170 shoes out of 3268). We took the mean of the rating and reviews columns for all of the shoes in the data frame.

```
best_shoesdf.mean()

Rating       3.242105
Reviews     40.551714
dtype: float64
```

If we only focus on the Rating and not the number of reviews, we get situations in which three people or nine people give a high review. This was seen in following screen shots. The data frames were coded to bring up the highest rated shoes in descending order. The screen shot captures all of the Nike and Adidas shoes with a perfect rating but there is a disparity in the number of reviews for Nike and Adidas.

Highest Rated Nike Shoes

	Product Name	Brand	Rating	Reviews
2635	Nike Air Max 90	Nike	5.0	9
2695	Nike Air Max 90	Nike	5.0	9
2736	Nike Air Force 1 '07 LV8	Nike	5.0	6
2965	Nike Air Zoom Wildhorse 5	Nike	5.0	6
3013	Nike Air Max 720 (OBJ)	Nike	5.0	5
3093	Nike Air Zoom Pegasus 36 Premium Rise	Nike	5.0	5
2966	Nike Air Max 720 Waves	Nike	5.0	4
2979	Nike Joyride CC	Nike	5.0	4
2661	Nike Air Force 1 '07 LV8	Nike	5.0	3
2721	Nike Daybreak SP	Nike	5.0	3
2737	Nike Air Max 90	Nike	5.0	3
2755	Nike Air Zoom Winflo 6	Nike	5.0	3
2768	Nike Air Force 1 '07 LV8	Nike	5.0	3
2885	Jordan Trunner NXT React	Nike	5.0	3
2890	Kyrie 5	Nike	5.0	3

Highest Rated Adidas Shoes

	Product Name	Brand	Rating	Reviews
267	Men's adidas Toe Side II Slippers	Adidas CORE / NEO	5.0	99
1800	Men's adidas Running Asweego Shoes	Adidas CORE / NEO	5.0	99
1849	MEN'S ADIDAS RUNNING NAYO 2.0 SHOES	Adidas CORE / NEO	5.0	97
1039	Men's adidas Sport Inspired Court Adapt Shoes	Adidas CORE / NEO	5.0	96
2505	Men's adidas Originals Liberty Cup Shoes	Adidas ORIGINALS	5.0	94
731	WOMEN'S ADIDAS ORIGINALS ARKYN PRIMEKNIT SHOES	Adidas ORIGINALS	5.0	92
1476	Men's adidas Originals Superstar MG Shoes	Adidas ORIGINALS	5.0	90
1049	Men's adidas Running Norad Shoes	Adidas CORE / NEO	5.0	89
1972	Men's adidas Sport Inspired Lite Racer RBN Shoes	Adidas CORE / NEO	5.0	89
523	Women's adidas Originals SL Andridge Shoes	Adidas ORIGINALS	5.0	79
2033	MEN'S ADIDAS ORIGINALS GAZELLE SHOES	Adidas ORIGINALS	5.0	79

Adidas shoes were reviewed at a much higher rate. There might be an issue in the sorting and the decision was made to establish a baseline for Ratings and Reviews that will give us a list of the top 170 shoes.

The means were used as a baseline and then we played with the rating and reviews by extracting the shoes in the 94th-95th percentile by pulling the shoes that had a Rating better than 4.55 and reviews greater 60.00. The shoes were then placed in a data frame.

```
] #creating a new data frame for shoes who have High ratings and High Reviews
  #Standard is set above the rating mean of 3.24 and reviews mean of 40.55
  ninety_fourth_df = best_shoesdf[(best_shoesdf['Rating']>4.55) & (best_shoesdf['Reviews'] > 60.00)]
  print(ninety_fourth_df)
```

The criteria filled the list with only two nike shoes and the rest were adidas. The screen shot below captures the two Nike shoes that met the criteria. (The list is included as one of the report csv files).

	Product Name	Brand	Rating	Reviews
2596	Men's Sport Inspired Roguera Shoes	Adidas CORE / NEO	5.0	78
2599	Kids-Unisex adidas Tennis Adizero Club Shoes	Adidas SPORT PERFORMANCE	4.9	84
2607	Men's adidas Tennis Wucht Indoor Shoes	Adidas SPORT PERFORMANCE	4.7	89
2750	Air Jordan 10 Retro	Nike	4.7	223
2780	Nike React Infinity Run Flyknit	Nike	4.6	68

CONCLUSION:
Given the data we can conclude that Adidas has the highest rated shoes and most reviewed shoes. There is a real passion from the Adidas customer to express their love for Adidas shoes.

In this particular data it appears that Adidas makes the better shoe however I'm not willing to give that as an answer to the final question. We do not have sale figures for Adidas and Nike. This data does not tell us how many shoes Nike and Adidas make and sell. We can assume that they make and sell shoes in the tens of thousands or hundreds of thousands. In order to say that Adidas makes better shoes we would need to see higher numbers in the Reviews column. The data resembles data from a store or a chain of stores instead of the world market. The data is not expansive enough to definitively state that Adidas or Nike make better shoes. For this particular store or regional market, we can say that the belief is that Adidas makes the better shoe, but other regions might have a different perspective.

Description of Output files:
Two output files.

First File: is a report on the mean of reviews, ratings and list of best shoes. The report is structured in columns that give a report statement and the outcome.

Second File: is a csv list of the highest rated and most reviewed shoes. The list is generated by shoes that meet a specified criteria that would place the shoes in the 94^{th} to 95^{th} percentile of the shoes in the entire data. Of note is the number of Nike shoes vs Adidas shoes in the list.