

Team 5 ETL Project- Shop Smarter

0

Table of Contents

- Team 5 ETL Project
- Purpose of the Analysis
- Key Documents
- Findings
- Coding Style
- API Calls
- Coding Documentation
- Authors
- Acknowledgments

Team 5 ETL Project- Shop Smarter

Team 5 chose a project that would examine the prices of a set list of items and compare these items across various zip codes across the United States. We would like to find if there are differences in prices based on location and look into the impact of product reviews. We have randomly selected five zipcodes in five different states (Alaska, Virginia, California, Florida, and Texas). A random list of items will be used to compare prices among the designated stores. The two primary considerations for shopper when considering a purchase are value and quality. Using data from the Walmart API, we will create a product information table, a product review table, a product review stat table and a trending products table. We will merge information from these various tables into one in order to create a data display that is easy to reference for the purpose of deciding which products to purchase.

Purpose of the Analysis

Our primary purpose is to gather data using the Walmart API to determine if the random items that were selected should be purchased by the consumer. Shoppers look for a good value (we will information on prices/sales) and quality (we will use information on consumer product reviews). This data will provide the pertinent information to the shopper when determining if they will make the purchase of the item.

Key Documents

The following data sets were used throughout the project.

Product Identification Table

0

Product Review Table

0

Overall Product Ratings Table

0

Megrged Product Information Table


0

Findings

- After doing research on prices across various zipcodes, we found that Walmart offers a 2 day ship to store meaning that no store has to have everything in stock. The products are available online and can be shipped to any other Walmart stores within 2 days, if needed. This makes the location of the store irrelevant as far as price comparison. Therefore, we shifted the focus on pulling data from three sources (product information, product reviews &)into one table. The merged table gives the shopper a comprehensive overview of the selected item. It indicates the price, the review score and the number of reviews for each item. As a shopper, this information is vital in terms of making a determination if you would like to buy a product.

Coding Style

Jupyter notebooks was used. The list of dependencies are as follows:

```
 import pandas as pd
import json
import requests
from sqlalchemy import create_engine
from config import wlm_api, myPass
from pandas.io.json import json_normalize
import time
```

API Calls

- Products API call: `http://api.walmartlabs.com/v1/items/12417832?apiKey={apiKey}&lsPublisherId={Your (http://api.walmartlabs.com/v1/items/12417832?apiKey={apiKey}&lsPublisherId={Your} LinkShare Publisher Id}&format=json`
- Product Review API call: `http://api.walmartlabs.com/v1/reviews/33093101?apiKey={apiKey}&lsPublisherId={Your (http://api.walmartlabs.com/v1/reviews/33093101?apiKey={apiKey}&lsPublisherId={Your} LinkShare Publisher Id}&format=json`
- Trending API call: `http://api.walmartlabs.com/v1/trends?apiKey={apiKey}&lsPublisherId=xyz&format=json` (`http://api.walmartlabs.com/v1/trends?apiKey={apiKey}&lsPublisherId=xyz&format=json`)

Authors

Lori Brady- Initial Work – SMU Data ETLTeam 5 (<https://github.com/loribeth18>)

Mabel Gutierrez- Initial Work – SMU Data ETLTeam 5 (<https://github.com/mabel912>)

Velindia Lucas – Initial Work – SMU Data ETLTeam 5 (<https://github.com/chele0630>)

Sisay Teketele – Initial Work – SMU Data ETLTeam 5 (<https://github.com/sisayyt>)

Ka-Ri Walker – Initial Work – SMU Data ETLTeam 5 (<https://github.com/ButtonWalker>)