Ranking Farmers Markets in Different Cities Based on MapReduce

Akhil Reddy Kommareddy
Department of Computer Science
Eastern Kentucky University
Richmond, Kentucky
akhilreddy_kommar@mymail.eku.edu

Stephen McQueen
Department of Computer Science
Eastern Kentucky University
Richmond, Kentucky
stephen_mcqueen15@mymail.eku.edu

Venkatesh Reddy Isanaka
Department of Computer Science
Eastern Kentucky University
Richmond, Kentucky
saivenkatesh_isan@mymail.eku.edu

***Abstract*— More than 10,000 farmer's markets have been registered in the US because of the rapidly expanding "farm-to-table" movement. Many of them have become a highlight for regular customers, with good quality food stalls and farm products that cannot be found in any other market. Farmers only make 15 cents of every dollar spent on the food they sell in a typical grocery store or supermarket. However, at a farmer's market, they keep all the money.**

Keywords—Hadoop, Map Reduce, Farmers Market.

I. Introduction

For the typical American household, purchasing organic, locally grown food is becoming more and more important. People are becoming more conscious of the foods they consume and where they come from. According to a survey, the demand for locally grown food has increased since the beginning of pandemic, so local farmers market is best to find high-quality products .
We are confident that every farmers market in United States is unique. Big market or small one, in major cities or small towns, farmers markets sustain their communities by helping local farmers, food makers, while we are trying to rank America's favorite markets, our intention is to uplift the role of farmers markets across the country, so in this project we are trying to find best nearby framer market

A great way to reconnect with your neighborhood, breathe in some fresh air, and be surrounded by nature's abundance of regional, organic, farm-to-table produce is to go down to your local farmer's market. Going through the aisles of the supermarket is far less appealing than taking in the sights, sounds, and colors of our neighborhood farmer's market. Supporting the men and women who toil in the fields to provide us with organic and delectably refreshing food is perhaps one of the finest reasons to frequent your local farmers markets. However, not all cities have farmers' markets. Through this effort, we can easily determine the finest places to go. This system's prepared to understand is survival. Farmers utilize improved farming methods to create healthy food for the native community, which in turn gives them the funds they need to survive. Each party benefits from the other's success in a cooperation that has evolved into a permanent arrangement. To safeguard the environment for present and future generations, organic production improves the health of the soil, plants, animals, and people.so to protect the farms and traditional farmers, people should start visiting the farmers markets with which they will also get benefited from organic products.

II. Literature Survey

The following will be a literature survey regarding the analysis of farmers markets. It will be summarizing several works that share the similar goal of either improving or searching for farmers markets through research possibly through quantitative means.

Farmers markets allow for and provide farmer communities with many social, economic, and ecological benefits (Local Food Communication and Marketing). Markets see increasing competition from several local food outlets, like for example community box shares, online marketplaces, and grocery stores. There exist many helpful practices for farmers market managers to reel in potential customers.

It is important for market managers to have knowledge and create promotional plans or techniques for their target customers. It is important to question why a customer wants to buy from and how they learned of their markets. Its managers cannot promote their markets individually; they need help from their vendors from their own market for promotion. For example, by allowing vendors to give their personal background, like for example their farm life, it can be used as a powerful selling point for farmers markets via their social media or farmer profiles. Additionally, managers can assist vendors in presenting their products to customers. Community groups can be valuable partners in market promotion. Examples include local businesses, associations, colleges, farm-to-table restaurants, and nonprofit organizations. Market managers use word of mouth as a strong method for local advertising of their farmers markets. Sharing promotion with local businesses or community organizations would significantly increase the chance of customers visiting the market. Social media is found to be one of the most valuable methods for contacting new and old customers. Facebook, for example, is commonly used among

marketing managers for not only attracting customers, but also recruiting volunteers and making public announcements. It is very important to establish clear goals during a planning phase; however, to ensure future success with marketing decisions, ongoing data collection and analysis is a fundamental requirement. Data related to this include customer surveys, feedback from venders, and vendor retention and EBT sales based on the specific season. [1]

Farmers markets have had a rise in popularity in many locations in the world. This surge requires an analysis of the social relations that these markets have with each other and their communities. Farmers markets are individually unique to each other; thus, it is important to understand healthy relationships between the market and local area. To achieve this, it is required to collect and analyze data from market managers, vendors, employees, and municipal employees working with non-profit farmers markets on their views as a stakeholder. Results verify that Tennessee farmers' markets, while their primary goal is to produce fresh food to a community, that the economic benefit and social standing define a successful market. [2]

Farmers markets are periodic, repeated activities when farmers and ranchers directly sell consumers a variety of fruits, vegetables, and other locally grown farm items. Markets restore and preserve local and regional food systems, having an abnormally large impact on it compared to their percentage of sales revenue. Farmers markets provide a broad range of beneficial impacts on communities, particularly economically, as shown by prior study. Inquiry into the social effects that markets have on communities has recently expanded in the United States. The above includes developing tools for producers as well as market stakeholders to measure their impact on producers and communities as well as increasing access to fresh food products and spreading awareness of the sustainable agricultural practices used mostly by producers. In order to identify current trends and to synthesize the available data describing the ways in which farm markets improve community well-being, this study evaluates prior research on the subject. It also identifies the areas which merit significant future investigation. Farmers markets offer numerous ways to influence the food system's ecological and health outcomes because of their close ties to both food producers and end users. [3]Farmers markets have a confirmed positive impact on local economies and farmers, The number of farmers markets in the US increased by an extra 86% over the ensuing ten years.

III. DATA SET

The data resource for the project is taken from the National Farmers Market Directory. It is maintained by AMS Marketing Services, is designed to provide members of the public with convenient access to information about U.S. farmers market. Market information included in the Directory is voluntary and self-reported to AMS by market managers, representatives from State farmers market agencies and associations, and other key market personnel.

The farmer's market dataset contains 59 columns. Broadly classifying columns as below for better understanding: [4]

FMID: is an identifier for each farmer's market and contains 7 digits number.
The market name is a free text column and allows special characters. [4]

Media Sites: Next five columns (Website, Facebook, Twitter, YouTube, other media) are supposed to contain information that will uniquely identify the market in one of the world wide web media sites. Most of the time this information is an URL to their page but there are no restrictions or validation so sometimes the value in these columns is free text. All five columns support blank or null values. [4]

Market Address: Next five columns (street, city, country, state, zip) is supposed to provide the address that will uniquely identify the market on the geographic map. All of them are free text columns, even the ZIP code column contains some special characters like '-'. However, one quality issue is that some of the names are lowercase, some capital case, and others have whitespaces before or after the string. This will cause issues if the user wants to do a direct string comparison. Also, a good way to avoid that and establish unification over these names of countries, cities, and streets can be for example to give each country a unique identifier and map the row that should belong to that country to the identifier instead. [4]

Season Date Time: The following eight columns containing the date and time for every season most likely represent the periods when the market is opened. The initial assessment clearly shows that these columns also are not consistent in terms of the format. There are occurrences where the period is stated as month A to month B and at the same time other occurrences where a period is more granular, for example from one date to another date. [4]

Location: The X and Y columns represent latitude and longitude, although the names of these columns are not meaningful enough, and the only way for a user to understand that is to see the column values. Another not-so-meaningful name is ' location'. The understanding is that it will contain map/geographic-specific information, however, it looks like it is more of a description of the place where the market is held. [4]

Market characteristic: The next 35 columns are Boolean columns containing Y-true and N-false for the given characteristic of a given market. An easily distinguishable

data issue is that sometimes they have a third or fourth value option like '-' or empty string which most likely indicates these tuples have neither false nor true values for these columns. [4]

Update Time: The last column (update time) representing the date and time when the record has been updated also does not follow a consistent Date Time format. There are records with only year, records with full date time, and records where month is present with a word instead of a digit. [4]

Data Pre-Processing: In data preprocessing first the columns which are not supposed to have null values is recognized. The columns are 'organic', 'baked goods', 'cheese', 'crafts', 'flowers', 'eggs', 'seafood', 'herbs', 'vegetables', 'honey', 'jams', 'maple', 'meat', 'nursery', 'nuts', 'plants', 'poultry', 'prepared', 'soap', 'trees', 'wine', 'coffee', 'beans', 'fruits', 'grains', 'juices', 'mushrooms', 'petfood', 'tofu', 'wild harvested' then null values in these columns are replaced with false. The rows which do not have city and state values are dropped alongside the 'organic' column has not the values in the true or false which is converted into true or false values. In the next step the categorical data of Website, Facebook, Twitter, and YouTube columns is replaced with the new Boolean variables then the columns are assigned with the new position values and at the final stage the columns which are much off for categorical values are dropped and the whole new data frame is transfer to the CSV file delimited by ';'.

IV. MEATHOD

Mapper: The mapper will take input from the preprocessed data file. Then it will loop over through every line, delimit the column values with the help of ';'. Next it will take city and state values as the key. It will check from the 9th to the 45th column if the corresponding value is true, then it will add up one to the rank variable. As we are having the digital media existence details from column 45th to column 47th we will give one extra credit for each kind of digital media platform the market is present. At the end, the mapper will send output in the format of a key and the count of the market which is 1 and they rank of the market.

Mapper Pseudo Code:

```
reader = csv.reader(sys.stdin, delimiter=';')
for line in reader:
 key = line[3]+','+line[5]
 rank = 0
for i in range(9, 45):
        if str(line[i]) == 'True':
            rank += 1
    for i in range(45,47):
        if str(line[i]) == 'True':
            rank += 2
  except:
    continue
  print(key,';',1,';',rank)
```

Reducer: The reducer will take the output from the mapper. In reducer, there will be a key-value pair defined earlier. Now the reducer will loop over the input from the mapper and it will pick the key, count, and rank of the particular record, and then it will check whether the key-value pair is present in the key-value pair object if it is not present then it will add it to the object otherwise it will take the previously existing instance of the object and it will do an aggregate count and aggregate rank of the particular key-value pair the aggregate count formula used for this particular use case is defined below and then end the reducer will print the key and the corresponding value which is present in the key-value pair which will be our final output.

Aggregate Rank =
(previous rank * previous count of markets) + current rank
/ (previous count of markets + 1)

Reducer Pseudo Code:

```
reader = csv.reader(sys.stdin, delimiter=';')

output = {} #dictonary to store reduced values.

for line in reader:
  key = line[0]
  count = line[1]
  rank = line[2]
  if output.get(key) == None:
    output[key] = str(count) + ',' + str(rank)
  else:
    oldValArr = output[key].split(',')
    oldCount = oldValArr[0]
    oldRank = oldValArr[1]

    newCount = int(oldCount) + int(count)
    newRank=
((int(oldRank)*int(oldCount))+int(rank))/int(newCount)

    output[key] = str(newCount) + ',' + str(int(newRank))

for key in output:
  print(key, ' ', output[key])
```

The Final output from reducer will be in key <count>,<rank>
format Ex: Fond du Lac,Wisconsin 2,25

V. RESULTS

In the proposed system we have preprocessed the data and then the mapper will take the preprocessed data set as input and output of the mapper is sending output in the format of a key and the count of the market which is 1 and they rank of the market. This output from the mapper will as input for reducer. In reducer, there will be a key-value pair that are predefined. Now the reducer will execute a loop over the input and it will pick the key, count, and rank that particular record, The reducer output will print the key and the

corresponding value which is present in the key-value pair. This will be our final output.

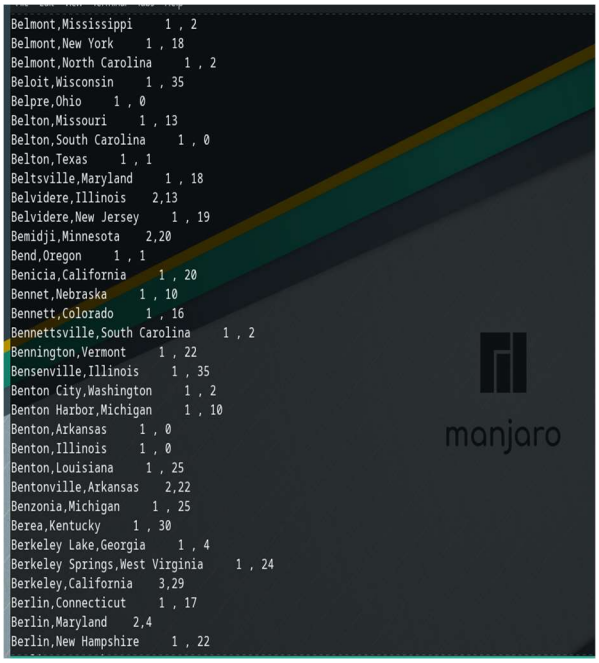

Fig 1. Sample Output

VI. CONCLUSION

Identifying customer needs for farmers' markets relies on the analysis of several variables which include product placements, product qualities, social standings with local area, and overall market planning. Aside from volunteers and vendors, customers are ultimately the lifeline in the economic success of farmers' markets; they are the foundation. One aspect that encourage customers new or old to visit a farmers' market is outside information regarding the quality of a market, whether it be an existence of an online platform for optimal communication or if a particular market possibly holds the produce they need. Ranking markets by their "market flexibility/availability" offers an optional consideration for potential customers when choosing which one they should be considering first. Additionally, data such as this can be potentially used for research purposes regarding correlation between farmers' market success and its market flexibility. Every farmers market in the US is unique. It may in big market or mini market, in major cities or small towns, farmers markets sustain their communities by helping local farmers, food makers, this can be achieved by ranking the framer markets in their area based on products available in market and based on social media presence for the market. So that the customers can decide which market to visit to get fresh market products, In the same way the framer markets also get more profits.

VII. FUTURE WORK

In the future the project results can be used to predict the markets that are best to find a particular type of product which is bestselling of that framer's market for example we can say that a particular market is famous for its oranges so that the customers required oranges can directly visit that market. The results can also be used as attributes in predicting the best city to live in. These results can be made on a website, and it can be shared in social media to gain more popularity. We can also add a customer review column to make the model more reliable by personal rating of the customers. similar goal of ranking farmers' markets should strive to add scoring weights to variables given research. For example, perhaps we consider a product like coffee being in the highest demand during a given year, then we should add a higher weight to the scoring to a market on whether it holds vendors selling coffee or not. This would enable for highly diverse data and scoring, thus allowing for potential methods like Top K's. Research and understanding of farmers' markets for these projects, however, should be essential for giving accurate rankings to each one.

VIII. REFERENCES

[1] "Farmers Market Research – Best Practices. Local Food Communication and Marketing.," [Online]. Available: https://fyi.extension.wisc.edu/localfoodmarketing/farmers-market-research/. [Accessed 3 December 2022].

[2] O. ". M. R. (. Tumlin, "Chancellor's Honors Program Projects.," [Online]. Available: https://trace.tennessee.edu/utk_chanhonoproj/2339.

[3] D. P. Warsaw, "The Economic, Social, and Environmental Impacts of Farmers Markets: Recent Evidence from the US," vol. 13, no. 6.

[4] [Online]. Available: https://github.com/pragyamis/DataCleaningProject_FarmerMarket.

[5] [Online]. Available: https://markets.farmland.org/.

[6] [Online]. Available: https://www.holidu.com/magazine/top-10-farmers-markets-us.

[7] [Online]. Available: https://www.gardensalive.com/best-cities-farmers-markets.