DATA ANALYSIS OF COMMODITY PRICES

A thesis submitted in partial fulfillment
of the requirements for the degree of

MASTER OF TECHNOLOGY

in

Computer Science & Engineering

by

KAPIL THAKKAR

Entry No. 2014MCS2124

Under the guidance of

Dr. Aaditeshwar Seth

Department of Computer Science and Engineering,
Indian Institute of Technology Delhi.

Jan 2016.

Certificate

This is to certify that the thesis titled **DATA ANALYSIS OF COMMODITY PRICES** being submitted by **KAPIL THAKKAR** for the award of **Master of Technology** in **Computer Science & Engineering** is a record of bona fide work carried out by him under my guidance and supervision at the **Department of Computer Science & Engineering**. The work presented in this thesis has not been submitted elsewhere either in part or full, for the award of any other degree or diploma.

Dr. Aaditeshwar Seth
Department of Computer Science and Engineering
Indian Institute of Technology, Delhi

Abstract

Supply demand imbalance, natural calamities etc. may not always be the reason behind the rise in the price of a commodity. It may be a result of artificial supply deficit planned intelligently by traders nexus to earn more profits through manipulation of supply of commodity and hence indirectly controlling their prices. Our attempt is to locate such hikes in prices which seem suspicious (we call them anomalies). We try to detect such anomalies by stating some hypothesis first and then algorithmically try to find the time-line during which it violates the stated hypothesis.

Acknowledgments

I would like to express my heartiest gratitude to our supervisors Dr. Aaditeshwar Seth for guiding this work with utmost interest, patience, care and scientific rigor. We thank them for setting high standards, giving us freedom to explore multiple facets of the problem and teaching us value of analytical thinking and hard work.

I would also like to thank Dipanjan Chakraborty who have helped a great deal by providing technical guidance and support whenever needed.

KAPIL THAKKAR

Contents

List of Figures

© *2015-16, Indian Institute of Technology Delhi*

List of Tables

Chapter 1

Introduction

1.1 Motivation

Supply demand imbalance, natural calamities etc. may not always be the reason behind the rise in the price of a commodity. It may be a consequence of artificial supply deficit planned intelligently by traders nexus for profiteering through manipulation of supply of commodity and hence indirectly controlling their prices. Our attempt is to locate such hikes in prices which seem suspicious (we call them anomalies). To detect and analyse the characteristics of anomalies in the prices of commodities. Currently we have considered the case of onion and based on that we are looking forward to develop library for any time-series.

1.2 Objective

Our objective is to create library with some set of functionalities to detect anomalies in the give input time series. Anomalies will be based on the hypothesis stated by the user. For particular scenario stated by user, execute appropriate algorithms and report the anomalies in the time series.

1.3 Relavence of Project

Anomaly detection techniques helps to explore many areas which are explained in detail in next chapter. If anomaly can not be justified by any natural parameter than it stats some serious issue. It states that something is wrong in the process and draws attention towards it. For example, unnecessary hike in the prices may be due to bad government policies, loopwholes in the system of commodity, etc. So our project will help journalist or end

© *2015-16, Indian Institute of Technology Delhi*

user interested in detecting abnormal behavior with the heklp of some hypothesis.

© *2015-16, Indian Institute of Technology Delhi*

Chapter 2

Literature Survey

2.1 What is Anomaly Detection?

According to wikipedia, anomaly detection (or outlier detection) is the identification of items, events or observations which do not conform to an expected pattern or other items in a dataset.

Here, our major focus is on detecting anomalies in the time series data. Time series usually considers data about price of some commodity, production, sells, etc. Usually, these time series follow some normal pattern. Some time series may be independent or behavior of some may depend on other time series. It may also consist of seasonality and trend along with some noise. So, considering all these factors our aim will be to detect some time line during which these time series does not follow a normal pattern.

Reasons for presence of anomaly may be different depending upon the type of time series. We have considered time series of onion data as a refernce. For that, reason for presence of anomalies are unseasonal rainfall, hoarding, price manipulation by traders' nexus, effect of import/export of onions in large amount, variation in production, etc. Being a seasonal crop, some part of onions are stored during large production season, so that when production is less in other season, demand can be met. But some traders hoards, keeping profit in mind. According to wikipedia, in economics, hoarding is the practice of obtaining and holding scarce resources, possibly so that they can be sold to customers for profit. So during this time also, we are able to see anomalies.

2.2 Onion Case

Onion is a staple ingredient for almost every Indian kitchen and hence its demand is almost constant throughout the year but not the supply. In order

© 2015-16, Indian Institute of Technology Delhi

to supply onions throughout the year, they are stored during harvest and released into markets in lean seasons. Its importance can be well estimated by the fact that it is one among few essential commodities and often rise in its price has resulted into downfall of state and central government.

One major tragedy in onion market occurred in end of year 2010. The prices of the onion increased so much that it was out of reach from poor people. There was a study conducted by the CCI (Competetive Commission of India) for this case and they created report on that [3]. In this [3], they tried to find out the reasons behind this scenario. They came with the following things in their study:

- Large wholesalers/traders mainly operates in metropolitan city markets and large number of farmers dispose their bulk of produce in nearby markets because of absence of storage facility, immediate cash need for loans, family expenses, purchase of inputs of next season, etc.
- Concentration of large storage capacities with traders,Vertical Integration of various market functions by onion traders(one name, many roles), Existence of established traders and barrier to new entry
- On December 23 of 2010 in The Times of India published in an article that on Tuesday alone, wholesale traders in Delhi bought onion at about Rs.34 per kg while it was sold in retail at Rs. 80 per kg, the margin of Rs. 46 per kg or 135 per cent.
- In the weeks of November and December, wholesale price remains high, so retailers do not get much profit, but even after that when wholesale price go down, retailers particularly in metro cities, show strong rigidity in holding price and earn margin from 60 to 110
- If we take this forward, then government policies also had a great role in the December 2010 high price episode (export of 1.33 lakh tones onion in October 2010).

So if we consider it overall, then yes, there was unseasonal rainfall in the month of September and October 2010, but after that also government policy

© *2015-16, Indian Institute of Technology Delhi*

regarding export of onion was unexplainable. The news article published in Times of India also questions why there is so much differnce in the wholesale and retail price of onions. Study also suggests that all the traders operating in the market have experience of many years (20 years on an average) and this is sort of family business too. Due to limited entries, there is also barrier for new trader to the market. So no new person enters and due to such large experience, they operate in the market together by forming the nexus. So many times they can alter the prices in many regions so that farmer has to pay money they decide. Traders have monopoly in onion markets and due that prices do not follow normal behavior of demand-supply and goes out of the way.

2.3 Other Cases

2.3.1 Sugarcane Case

This [4] study on sugarcane shows the connections between Politicians and Sugar Mills in Maharashtra and explains how such connection may benefit to firms and politicians. Sugar mills are cooperatives and regions are formed according to mills present. Each sugarcane farmer has to sell his produce to the mill present in his region, he can not sell it to some other mill. Each mill pays its farmers a single price per metric tonne of cane every year, based on weight (not on the quality).

This study investigates how price of the sugarcane, paid to farmers, changes in the election year. Usually, chairman of the sugar mills are politicians who stands in election. They need funding for elections, so here author explains how sugarcane mills and election funding may be related. And if some chair person wins the elction then what is effect on prices. Some findings are:

- Prices are lower by about Rs. 20 a ton in politically controlled mills during election years
- The results are robust to including rainfall and mill capacity as controls, as well as including mill-specific outcomes such as the recovery rate

© 2015-16, Indian Institute of Technology Delhi

- sugar produced per unit cane, a measure of productivity - as well as various other mill level shocks such as mill breakdowns and cane shortages

- Price fall may be due to mill closure, i.e. mill is not operating profitably, but politicians has kept mill open as a way to garner votes. But analysis shows that mill closure is not affected by political control

- Paying farmers Rs. 20 per ton less for their cane amounts to a total of Rs. 6 million

- Mills whose chairmen won national elections pay Rs. 80 per ton more in the year after elections

- Author finds that when the party affiliated with the mill chairmen is in power in Maharashtra, the mill pays Rs. 23 more in cane price and also Chairmen who win national elections seem to be able to keep their mills open far more successfully than chairmen who lose

So here author has strong belief that all facts are pointing that funding for election campaign comes from these sugarcane mills if mill is politically controlled. Reason why farmers suports this may be that, with average probability 1/3rd of winning election, on an average farmer gets Rs. 27 on their principal of Rs. 20, so still in profit. The overall effect on farmer welfare is difficult to determine. On average, cane prices and recovery rates in politically connected mills are no different than those in non-politically connected mills, and the levels of public goods are no different either.

So this example explains how time series may go out of their normal behavior though there is no supply-demand crisis or any other case.

© *2015-16, Indian Institute of Technology Delhi*

Chapter 3

Study of Onion Data: Collection and Analysis

3.1 System

For now, we are only working over onion data. We have three actors in model: Farmers who are producers of onion, traders who are collectively responsible for supply of onions across countries and consumers who purchase onions.

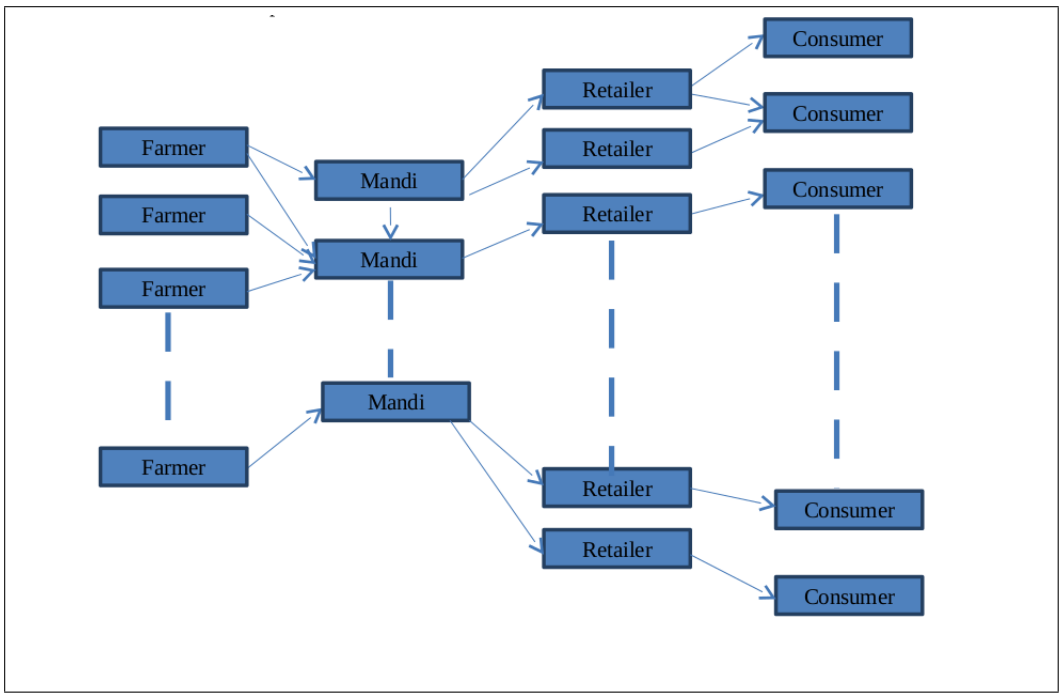

Figure 3.1: Normal Supply Chain

Farmers sell their produce to traders in nearest mandi offering better price. These traders sell these commodities to traders in other mandis or to retailers. Consumers purchase commodities for consumption from one of the retail stores. This way commodity reaches consumers from farmers following a huge chain of traders and retailers.

© 2015-16, Indian Institute of Technology Delhi

Under APMC act, mandis were established at different places across country so that farmers can sell their produce directly in mandi and get good returns (wholesale price). There are around 1500 mandis located in different places across country which log their daily arrival of onion, minimum, maximum, modal selling price per quintal of onion data to AGMARKNET. Retailers purchase from these mandis and sell to end customers at retail price. There are around 70+ centres across country which maintains retail price of onion on Ministry of consumer affairs website.

3.2 Data we have

We have following data:

1. Daily wholesale price of onion for 1514 mandis
2. Daily arrival of onion information for 1514 mandis
3. Daily retail price of onion for 76 centres
4. Dates and location for hoarding reports from news articles
5. Longitude and latitude of mandis and centres

Onion Data was collected from the the government websites,[1] for arrival and wholesale price data and [2] (Department of Consumer Affairs) for retail price data. Crawlers were written to collect the data from these datewise for the period of approximately 9.5 years, starting from Jan 2006 to Jun 2015. More data can be added simply by running crawlers again.

3.3 Normal market behaviour

1. Wholesale price is inversely proportional to arrival of commodity higher production of crop will lead to more and more crop hitting market for sell. Hence more arrival which will result in surplus supply leading to drop in wholesale prices.

© 2015-16, Indian Institute of Technology Delhi

2. Retail price is directly proportional to wholesale price commodities reach customers through a long chain of traders and retailers, adding value at every stage of chain. So, retail price at which customers purchase commodities are more than wholesale.

Any divergence from these characteristics of normal market leads to suspicious price hike situations/anomalies.

3.4 What are the reasons for anomaly?

Primarily there are 3 main reasons of anomaly.

1. **Government Policies:** When the price is low in the country, still government allows the export of onion in large amount, or supports it by keeping low minimum price then the prices can rise up drastically.

2. **Unseasonal Rainfall:** Due to insufficient rainfall, heavy rainfall or unseasonal rainfall, onion crop may get affected and the produce is low and wholesale price may rise up. But, this reason still is validating that wholesale price is inversely proportional to arrival, it may be just prices will be little higher than it was supposed to be.

3. **Hoarding:** When traders/wholesalers store the onion and does not release the stock in the market in the expectation of the good prices in the future, it will create the artificial deficit in the market and will shoot up the onion prices in the retail market due to low arrival in the retail market. The reason people do this is to expect the higher prices in the time of low production or may be for security. For example, if it is expected that this year the rainfall will be very low, then people may think that, due to that, production will be low in the future and so they will start storing onion right now and that will also create deficit in the market and price will go up.

So our study will focus on detection of anomalies in data and if possible comment on the possible reason for the anomaly.

© *2015-16, Indian Institute of Technology Delhi*

3.5 Mapping of wholesale price to retail price

Voronoi Diagram is used to map every mandi to nearest possible centre. The centres with retail data were considered fixed points and country was divided into 76 regions. All the mandis falling in that region are mapped to the respective centre.

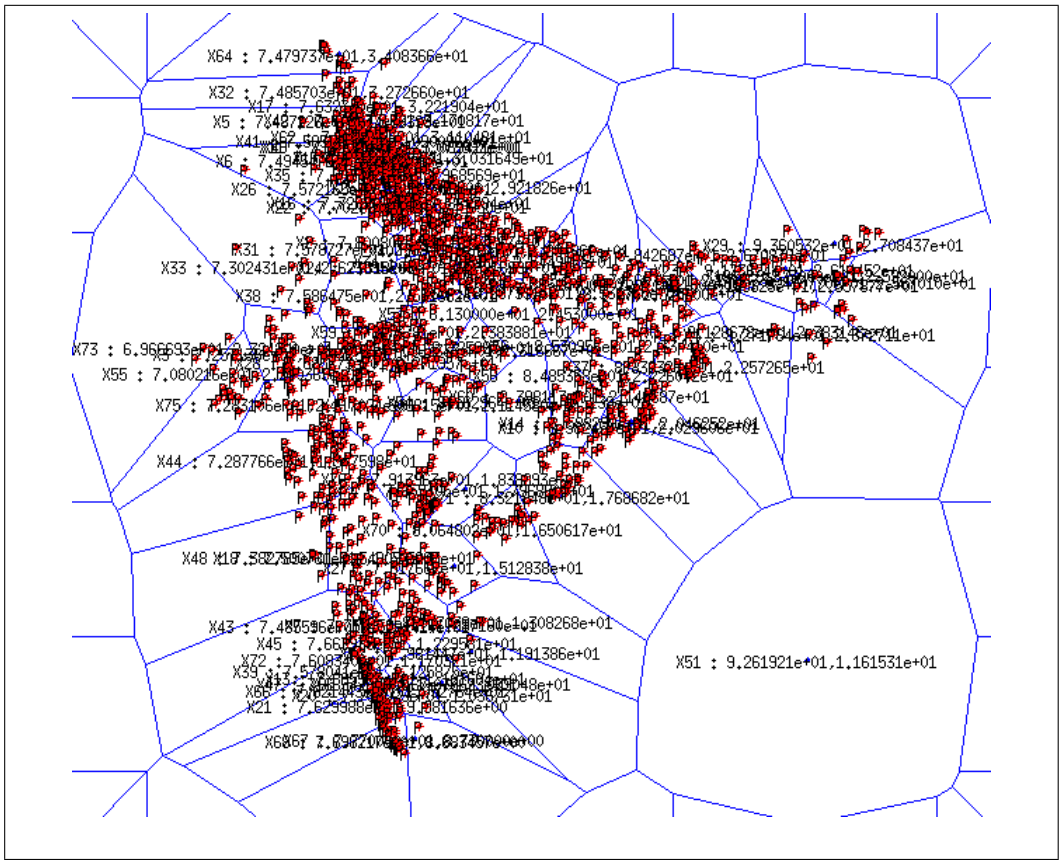

Figure 3.2: Voronoi Diagram

After all the mandis are mapped to their nearest centre, wholesale and arrival at every centre is computed. Wholesale price at centre is average of modal price of all mandis in its region and arrival was computed as the sum of the arrival at the mandis in its region. While calculating the wholesale price, the distance between centre and the mandi was not considered.

Corresponding to every date and location of hoarding news report, values like current year arrival, last year arrival, Percentage difference in wholesale-retail etc. were computed. Following table resulted from these computations.

© 2015-16, Indian Institute of Technology Delhi

3.6 How to Define Anomaly?

© 2015-16, Indian Institute of Technology Delhi

Chapter 4

CHAPTER NAME

4.1 SECTION NAME

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

4.2 SECTION NAME

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

© *2015-16, Indian Institute of Technology Delhi*

4.3 SECTION NAME

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

© *2015-16, Indian Institute of Technology Delhi*

Chapter 5

Conclusion

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

© *2015-16, Indian Institute of Technology Delhi*

Bibliography

[1] Agriculture marketing. http://agmarknet.dac.gov.in/. (Visited on 01/04/2016).

[2] fcainfoweb.nic.in/pmsver2/reports/report_menu_web.aspx. http://fcainfoweb.nic.in/PMSver2/Reports/Report_Menu_web.aspx. (Visited on 01/04/2016).

[3] PG Chengappa, AV Manjunatha, Vikas Dimble, and Khalil Shah. Competitive assessment of onion markets in india. *Institute for Social and Economic Change. Competition commission of India*, 1:86, 2012.

[4] Sandip Sukhtankar. Sweetening the deal? political connections and sugar mills in india. Technical report, Mimeo, 2008.

© *2015-16, Indian Institute of Technology Delhi*