

Impact of sanctions on food prices in Iran

rahmani_pour.payam@stud.hs-fresenius.de

HS-Fresenius: Data Science for Business

Abstract

Unprecedented food market inflation has happened due to the United States and the United Nations imposing sanctions on Iran, which will undoubtedly impact the Iranian people's food security. The current research examines the effect of sanctions on Iranian food prices from 2012 until 2022. The impact of sanctions on the retail price of food in Iran was evaluated using time series analysis. The groupings of pulses, nuts, meat, fish, eggs, cereals, and tubers had the most remarkable inflation rate. High inflation in food prices makes it difficult for most individuals to keep a typical healthy diet. Iranians may soon be more prone to chronic diseases due to this.

Contents

1	Introduction	2
2	Data Preparation & Overview	6
2.1	Data Cleaning & Processing	7
2.2	Data Analysis	15
3	Discussion	25
4	Conclusion	26
5	Limitations and Challenges	26
	References	27

Rendered at 06 February, 2023

Word count: 5481

1 Introduction

Several nations, most notably the United States and international organizations, have imposed sanctions on Iran. Prior to Russia overtaking Iran as the nation with the highest sanctions after the latter invaded the adjacent Ukrainian territory in February 2022 (Bloomberg 2022). Since 1979, when the American embassy in Tehran was taken over, the United States has restricted trade and other activity with Iran under a number of legal pretexts (Josh Levs 2012). After the hostages were freed in January 1981, these sanctions were suspended; however, the United States reinstated them in 1987 in reaction to Iran's activities from 1981 to 1987 in opposition to American and foreign military vessels in the Persian Gulf as evidence of its support for terrorism (Wikipedia 2022). In 1995, the restrictions were extended to include companies doing business with the Iranian government. (Zirulnick 2011) After Iran declined to abide by UN Security Council Resolution 1696, which ordered Iran to cease its uranium enrichment program, the third sanction was put in place in December 2006 by UN Security Council Resolution 1737. American sanctions first targeted the export of refined petroleum products, investments in the oil, gas, and petrochemical sectors, and dealings with the Islamic Revolutionary Guard Corps (IRGC). It included shipping, web hosting for business ventures, domain name registration, and Finance and insurance operations (ecfr.gov 2010). Subsequent UN resolutions have increased sanctions on Iran.

Iran's economy and people have suffered greatly due to sanctions. Since 1979, the United States has been at the forefront of initiatives made by nations to use sanctions to change Iran's policies, notably its program for uranium enrichment, which Western nations worry is meant to create the capacity to manufacture nuclear weapons. Iran argues that its nuclear program is only used for peaceful reasons, such as producing power and treating patients (Nichols 2012).

United States senators used the the failure of the nuclear talks between Iran and the West to justify tightening economic sanctions against Iran (Lakshmanan 2013). During a meeting in Lausanne, the P5+1 and Iran reached a provisional agreement on a framework on 2015-04-02. Once it is finalized and implemented, the framework will lift most of the sanctions in exchange for Iran's nuclear program being constrained for at least ten years. The Joint Comprehensive Plan of Action, the final accord, was approved on October 18, 2015 (JCPOA 2019). Consequently, on January 16, 2016, UN sanctions were withdrawn. Donald Trump, the president of the United States, said that the country would leave the Iran nuclear agreement on May 8, 2018. The United States reinstituted sanctions in November 2018 and extended them to include Iran's banking sector from 2019 till now. The Financial Action Task Force added Iran to its blacklist on February 21, 2020.

The sanctions imposed by the United States on Iran have had detrimental humanitarian effects on Iranian society. Iranians' rights to food, health, education, and other human rights have suffered due to the sanctions' broad scope. Following the US sanctions against Iran, unprecedented inflation has also occurred in Iran, affecting the food market.

Definition of Data science. Data science definition will assist in clarifying the unique characteristics of data science projects and outline the suggested management style in this study. There is general agreement among the writers that combine data science with well-

established branches of research about the domains that support and expand the data science tree. For example, (Granville 2014) describes data science as the data mining, machine learning, operations research, six sigma, automation, and domain knowledge at the convergence of computer science, business engineering, while (Martinez, Viles, and Olaizola 2021) claims that data science is a multidisciplinary intersection of mathematic expertise, business acumen, and hacking skills. According to (Loukide 2011), data science requires understanding conventional computer science, mathematics, and the arts. (Conway 2023) also provides a Venn diagram in which data science is represented as the union of a) hacking abilities, b) math and statistics knowledge, and c) substantive expertise. In contrast, many challenges in data science are statistical engineering problems with more extensive, more complicated data that may call for distributed computing and machine learning techniques in addition to statistical modelling, according to the (Martinez, Viles, and Olaizola 2021).

To comprehend the function of data science in business & industry sectors and its potential domain applications, only a few writers have examined the underlying goal of data science (Martinez, Viles, and Olaizola 2021). In contrast to (Das et al. 2015), according to which “statistical and machine learning approaches on huge multi-structured data in a distributed computing environment to detect correlations and causal linkages, classify and forecast events, identify trends and insights, and assess probability, interest, and attitude,” The authors of (Jeff Saltz et al. 2018) define data science as the analysis of data to solve problems and generate insights. For them, data science integrates knowledge of statistics, data management, and software development. According to (Jeffrey Saltz, Armour, and Sharda 2018), the study of computational concepts, techniques, and systems for obtaining and organizing information from data is referred to as data science. To better comprehend its place among conventional employment roles, there has also been an increase in interest in describing the work done by data scientists and outlining the abilities required to become one. As (Warden 2011) presents a “unicorn” perspective of data scientists and claims that data scientists handle everything from finding data to processing it at scale, visualizing it, and writing up as a story, (Wills 2012) defines a data scientist as someone who is “better at statistics than any software engineer and better at software engineering than any statistician”.

Among the thousands of potential definitions for “data science,” (Martinez, Viles, and Olaizola 2021) uses the remarks from above as a guide and offers the following new definition: “Data science is an multidisciplinary field that lies between computer science, mathematics and statistics, and comprises the use of scientific methods and techniques, to extract knowledge and value from large amounts of structured and/or unstructured data”. Therefore, (Martinez, Viles, and Olaizola 2021) concludes from this definition that data science programs strive to resolve complicated real-world issues using data-driven methodologies. In this way, data science applies to practically all of the current industries and fields: Financial services (customer segmentation, risk analysis, algorithmic trading), banking (fraud detection, credit risk modelling, customer lifetime value), healthcare (medical image analysis, drug discovery, bio-informatics), manufacturing optimization (failure prediction, maintenance scheduling, anomaly detection), e-commerce (targeted advertising, product recommendation, sentiment analysis), and transportation. Although data science may be thought of as an application-neutral profession, (Martinez, Viles, and Olaizola 2021) think that it is strongly advised to know the application domain in order to extract value from data

properly.

Research Methodology. (Martinez, Viles, and Olaizola 2021) presented a critical literature review based on the initial concepts and a comparison of literature on data science project management. They gathered information on the use of methods for data science from a variety of sources, including books, websites, papers, and open internet publications platforms. Hereunder a list of these methodologies with a summary of definitions is gathered.

Cross-Industry Standard Process for Data Mining (CRISP-DM) (Shearer 2000) is a 1996 open standard process model that outlines typical techniques used by data mining professionals. SPSS and Teradata created it. It offers an iterative process that is organized, clear, and well-documented. CRISP-DM divides a data mining project’s lifespan into six parts: business understanding, data understanding, data preparation, modeling, assessment, and deployment.

Microsoft Team Data Science Process (TDSP) by Microsoft (Grady 2016) is a “data science technique that is agile and iterative and helps to improve team communication and learning.” It has a wide range of tools and utilities that make usage more manageable and is well documented. Unfortunately, TDSP relies highly on Microsoft services and rules, making wider usage difficult.

In a 2017 white paper, Domino Data Lab described the lifecycle of a data science project (Lab 2017). It adopts “a comprehensive method for handling every stage of a project’s lifetime, from planning through execution and monitoring.” It was motivated by CRISP-DM, agile, and some customer insights. Three tenets serve as the foundation for the methodology: a) “Expect and welcome iteration,” but “avoid iterations from materially postponing projects or diverting them from the immediate objective.” b) By developing components that may be used in other projects, you can “enable compounding cooperation.” c) “Anticipate auditability demands,” and “preserve all relevant artifacts related to a model’s development and deployment.”

RAMSYS by (Moyle and Jorge 2001) is a way to facilitate remote, quick-team data mining operations. It is designed for dispersed teams and is guided by the following principles: light management, starting at any time, stopping at any moment, flexibility in problem-solving, knowledge exchange, and security.

Agile Data Science Lifecycle by (Jurney 2017) presents a framework for doing data science paired with agile principles. According to this methodology, a web application is the most efficient and appropriate way for data science to benefit organizations. As a result, from this perspective, doing data science evolves into developing applications that describe the steps involved in applied research, such as rapid prototyping, exploratory data analysis, interactive visualization, and applied machine learning.

(Crowston et al. 2019) establishes a theoretical framework for socio-technical affordances for stigmergic cooperation. This technique aims to improve coordination in data science teams by using lessons learned about coordination from the development of free/libre open source software (FLOSS).

Development Workflows for Data Scientists by Github and O’Reilly Media assembles a variety of procedures and best practices for data scientists (Byrne 2017). The article looks at how

several data-driven businesses are enhancing their data science development processes. The suggested data science methodology has an iterative design: Asking a compelling question, looking at earlier work, gathering data, exploring it, modeling it, testing it, documenting it, deploying it to production, and communicating the findings are all steps in the process.

Big data ideation, assessment and implementation by (Vanauer, Bohle, and Hellingrath 2015) is an approach to direct the production, evaluation, and management of big data ideas. It is based on corporate architectural management, IT value theory, workgroup ideation processes, and the big data 4Vs (volume, variety, velocity, value, and veracity).

By connecting business objectives with technological execution, (Kaufmann 2019) Big Data Management Canvas is a reference model for big data management that operationalizes value generation from data.

(Larson and Chang 2016) provide a framework built on the fusion of business intelligence (BI), quick analytics, and data science with agile concepts. The strategic tasks are divided into two levels: Fast analytics and data science are included in the bottom layer, whereas BI delivery is included in the top layer.

Systematic Research on Big Data by (Das et al. 2015) investigate the formalization of methods for data-driven research. Eight Agile analytic phases comprise the procedure described here; they begin with a given dataset and finish with the research result. In this regard, (Collier 2012) suggested development style for the process. Data is first extracted and cleansed. Although the data form the basis of data-driven research, the authors contend that the process should begin with a consideration of the study's goals and the data. The preliminary data analysis, in which specific patterns and information are revealed from the data, is the subsequent step. Unsupervised techniques may be utilized to extract these patterns, which will be used to determine the research's goal and entry point. The definition of the study objective or research hypothesis follows from the produced data patterns and findings. More work on the retrieved data may be done after the configuration of the study aim has been determined. If not, the study objective must be changed, and early data analysis procedures like descriptive analytics must be repeated. Therefore, given the purpose of the study, further analysis can be done by choosing the most relevant characteristics and creating machine learning models. Further analysis is done on the results produced by these models or prediction systems. Iterative improvements may be made to the output assessment inside an agile process. Last but not least, it is critical to effectively convey and publish the outcomes utilizing insightful infographics and data visualization tools. It is possible to carry doing these processes repeatedly until the desired outcomes or level of performance is attained.

This technique recommends starting the workflow described above with a small portion of the dataset, identifying problems as they arise, and only expanding to the complete dataset once everything is working as it should. In addition to using iterative Agile planning and execution, a generic dataset and standardized data processing may make data-driven research more systematic and consistent. Generally, this framework offers a procedure for doing methodical research on big data and is backed by Agile development approaches. Agile may encourage regular communication and value-driven development in a gradual, iterative way. However, the writers need to spend more time outlining the responsibilities and interactions of the team members throughout the various phases of the project. Since data science is

situated between research activities and commercial applications, it is beneficial to have a technique that can extract the most excellent qualities from at least one of these worlds. Overall, this approach aligns more with the research and academic world. By comparing all of these methods, this method is more suited for this research, and the methodology of systematic research will use for this context.

Structure. There are three primary chapters in this data analytics project that follows. The situation and the project's purpose will comply in this first part. The second section is a thorough description of all crucial procedures taken during the whole analytical process. The last section is a report for an international body that summarizes the study and answers the inquiry.

Analysis goal. The current research aims to ascertain if the sanctions have affected Iranian family food costs and which category sees the most rise. To achieve this goal as effectively as possible, the relationship between the date of imposing sanctions and the food costs rise would be investigated.

2 Data Preparation & Overview

The datasets that will be used in this project has been made available by World Food Programme Price Database. under a public license (humdata 2022), and to access the data source visit sources (Database 2022).

Setup and Data Import: When opening R Studio first the libraries `lubridate`, `hms`, `ggplot2` and `tidyverse` and its sub-libraries, which contain important functions for processing, analyzing and visualizing data, must be loaded.

```
library(tidyverse)
```

```
library(ggplot2)
library(hms)
library(lubridate)
```

Next, the CSV file will be imported as a data frame from the website.

```
foodprices <- read.csv(url("https://data.humdata.org/dataset/6df76343-1bd9-488a-af3c-1e"))
```

Dataset Overview: Now by the `str()` function can learn more about the data structure of our table. Each sample in the data sets contains details on the pricing of a specific product. There are 762 items in the table, each with 14 variables.

```
str(foodprices)
```

```
## 'data.frame':      872 obs. of  14 variables:
## $ date      : chr  "#date" "2012-04-15" "2012-04-15" "2012-04-15" ...
## $ admin1    : chr  "#adm1+name" "Tehran" "Tehran" "Tehran" ...
## $ admin2    : chr  "#adm2+name" "Tehran" "Tehran" "Tehran" ...
## $ market    : chr  "#loc+market+name" "Tehran Market" "Tehran Market" "Tehran Market" ...
## $ latitude   : chr  "#geo+lat" "35.72" "35.72" "35.72" ...
## $ longitude  : chr  "#geo+lon" "51.4" "51.4" "51.4" ...
## $ category   : chr  "#item+type" "cereals and tubers" "meat, fish and eggs" "miscellaneous" ...
## $ commodity  : chr  "#item+name" "Rice" "Eggs" "Sugar (white)" ...
## $ unit       : chr  "#item+unit" "KG" "1.8 KG" "KG" ...
## $ priceflag  : chr  "#item+price+flag" "actual" "actual" "actual" ...
## $ pricetype  : chr  "#item+price+type" "Retail" "Retail" "Retail" ...
## $ currency   : chr  "#currency" "IRR" "IRR" "IRR" ...
## $ price      : chr  "#value" "26860.0" "52760.0" "15890.0" ...
## $ usdprice   : chr  "#value+usd" "2.1909" "4.3034" "1.2961" ...
```

Here we can see that there are 14 different columns. The first column shows the date from the second to fifth columns about the store's location in Tehran. The next one is about the category of items and the name of the items. Furthermore, we have the items' units, price flags, and price types. The last three columns are about currency, price in IRR, and price in US dollars.

2.1 Data Cleaning & Processing

At first, we knew that the second row clarified our columns. Now we could remove this row.

```
foodprices <- foodprices[-c(1), ]
```

Checking for Missing Values:

The `is.na()` routines will be used first to identify **NA** entries, which will then be summarized for each variable separately. The second phase will show the number of fields with **NULL** values.

```
# Printing amount of missing values for each of the i = [1:14] features
for (i in 1:14) {
  message(sum(is.na(foodprices[i])), " missing values in column ", i)
}
```

```
## 0 missing values in column 1
```

```
## 0 missing values in column 2
```

```
## 0 missing values in column 3
```

```
## 0 missing values in column 4

## 0 missing values in column 5

## 0 missing values in column 6

## 0 missing values in column 7

## 0 missing values in column 8

## 0 missing values in column 9

## 0 missing values in column 10

## 0 missing values in column 11

## 0 missing values in column 12

## 0 missing values in column 13

## 0 missing values in column 14
```

It is shown that there is not any missing value in dataset.

It is time to check the price's column for any zero value and omit them.

```
foodprices$price <- as.integer(foodprices$price)
foodprices[foodprices==0] <- NA
foodprices <- foodprices[complete.cases(foodprices),]
```

Understandably, there were 13 rows with zero values that were omitted. Checking **price** for minimum, average and maximum values.

Exploring Distinct Values: The method *n_distinct()* is used to acquire a summary of how many distinct values each feature can store. It gives the total number of unique values in a column.

```
library(dplyr)
```

```
# Printing distinct values for each of the i = [1:12] features
for (i in 1:14) {
  message(n_distinct(foodprices[i]), " distinct values in column ", i)
}
```



```

## 114 distinct values in column 1

## 1 distinct values in column 2

## 1 distinct values in column 3

## 1 distinct values in column 4

## 1 distinct values in column 5

## 1 distinct values in column 6

## 8 distinct values in column 7

## 34 distinct values in column 8

## 15 distinct values in column 9

## 1 distinct values in column 10

## 1 distinct values in column 11

## 1 distinct values in column 12

## 453 distinct values in column 13

## 620 distinct values in column 14

```

Variables that can already be validated as cleaned:

- admin1, admin2, market, latitude and longitude attributes are the same for every item
- pricetype and currency attributes are the same for every item
- The Priceflag attribute has two distinct values, and by exploring the table, it is evident that only ten rows have a forecasting flag

Questionable variables:

Whether there are nonsense or zero variables for the prices should be checked. Also, there is a discrepancy between **IRR Prices (444/761)** and **USD Prices (583/761)**. One scenario could be that the price is stable in the home currency, but the exchange rate has changed.

Moreover, it can be seen that there are eight different categories and 37 different items. However, some lower or upper case conditions and white spaces should be checked to determine if there is any duplicate. Furthermore, 16 different scale units must be checked to see if every item has the same unit in the time series.

```
# Declare trim function to delete leading and trailing white spaces
trim <- function(x) gsub("^\\s+|\\s+$", "", x)
# Trimming column $category and $commodity
foodprices$category <- trim(foodprices$category)
foodprices$commodity <- trim(foodprices$commodity)
```

```
message(n_distinct(foodprices[7]),
        " distinct values in column category after trimming")
```

```
## 8 distinct values in column category after trimming
```

```
message(n_distinct(foodprices[8]),
        " distinct values in column commodity after trimming")
```

```
## 34 distinct values in column commodity after trimming
```

```
# Converting all character of data frame to lowercase
foodprices <- mutate_all(foodprices, .funs=tolower)
```

```
message(n_distinct(foodprices[7]),
        " distinct values in column category after conversion to lowercase")
```

```
## 8 distinct values in column category after conversion to lowercase
```

```
message(n_distinct(foodprices[8]),
        " distinct values in column commodity after conversion to lowercase")
```

```
## 34 distinct values in column commodity after conversion to lowercase
```

After trimming and in the conversion to only lowercase, no values have been excluded from each of the two variables **category(8/761)** and **commodity(37/761)**.

By searching in the dataset, it could find out that the commodity and category until 2021 are stable, but for 2022 we have more items. Because of this, it would be better if we split our investigation because of consistency and just analysis until last 2021.

```
message(n_distinct(foodprices[foodprices$date<"2022-01-01",7]),
        " distinct values in column category until 2022")
```

```
## 5 distinct values in column category until 2022
```

```
message(n_distinct(foodprices[foodprices$date<"2022-01-01",8]),
       " distinct values in column comodity until 2022")
```

```
## 5 distinct values in column comodity until 2022
```

There are five different items with five different categories.

For the simplicity of this project the **admin1**, **admin2**,**market**, **latitude**, **longitude**, **priceflag**, **pricetype** and **currency** variables are simply dropped from the data set.

```
foodprices <- foodprices[, -c(2:6,10:12)]
```

Identifying if commodities have the same unit along time series. Firstly knowing our unique commodities and units would be helpful.

```
unique(foodprices[foodprices$date<"2022-01-01",3])
```

```
## [1] "rice"           "eggs"           "sugar (white)"  "oil (vegetable)"
## [5] "lentils (green)"
```

```
unique(foodprices[foodprices$date<"2022-01-01",4])
```

```
## [1] "kg"      "1.8 kg" "810 g"  "900 g"
```

To check the equality of variables.

```
library(stringr)
foodprices[foodprices$date <"2022-01-01",] %>%
  group_by(item = str_extract(commodity,"^\\w+(?=\\s*)")) %>%
  summarise(isUnitSame = n_distinct(str_extract(unit,"[a-z]+$"))==1)
```

```
## # A tibble: 5 x 2
##   item      isUnitSame
##   <chr>    <lgl>
## 1 eggs     TRUE
## 2 lentils  TRUE
## 3 oil      TRUE
## 4 rice     TRUE
## 5 sugar    TRUE
```

For sure, all commodities have the same unit from 2012 to 2021.

Transforming Data Types:

All column data formats should be acceptable for their context to evaluate the data correctly. In the case of this data, the majority of the variables need to be changed.

```
# Changing data types
foodprices$date <- as.Date(foodprices$date)
foodprices$category <- as.factor(foodprices$category)
foodprices$commodity <- as.factor(foodprices$commodity)
foodprices$unit <- as.factor(foodprices$unit)
foodprices$price <- as.integer(foodprices$price)
foodprices$usdprice <- as.numeric(foodprices$usdprice)
```

Next new variables **day** (classes: Mon - Sun), **month** (Jan - Dec) and **year** are created according to the date frame of the date variable.

```
### Create $day and $month and $year
foodprices$day <- weekdays(as.Date(foodprices$date))
foodprices$month <- months(as.Date(foodprices$date))
foodprices$year <- year(as.Date(foodprices$date))
```

Converting these new variables to proper characteristics.

```
# Convert to factor and integer
foodprices$day <- as.factor(foodprices$day)
foodprices$month <- as.factor(foodprices$month)
foodprices$year <- as.integer(foodprices$year)
```

Exploring Data Consistency:

```
## Minimum Price:2691
```

```
## Average Price:90742.5366336634
```

```
## Maximum Price:594000
```

The minimum,average,maximum values are realistic and consistent.

Go deep and zoom into each commodity to find the minimum price.

Table 1: Min Price For Each Item

commodity	min(price)
eggs	6085
lentils (green)	2691
oil (vegetable)	25790
rice	26860
sugar (white)	15800

It was evident that min price for two items of **eggs (6085 IRR)** and **lentils (2691 IRR)** is unrealistic.

These prices should be manipulated to solve this problem..Firstly these prices should be identified.

```
foodprices %>%
  filter_all(any_vars(. %in% c(2691,6085,29910)))
```

```
##           date           category      commodity  unit price usdprice
## 1 2020-02-15 meat, fish and eggs          eggs 1.8 kg  6085    0.1449
## 2 2020-04-15 meat, fish and eggs          eggs 1.8 kg  6085    0.1449
## 3 2020-06-15 meat, fish and eggs          eggs 1.8 kg 29910    0.7121
## 4 2020-06-15      pulses and nuts lentils (green) 900 g  2691    0.0641
##           day      month year
## 1  Saturday February 2020
## 2 Wednesday   April 2020
## 3   Monday     June 2020
## 4   Monday     June 2020
```

Changing the prices with the new value.

```
foodprices['price'][foodprices['price'] == 2691] <- 269100
foodprices['price'][foodprices['price'] == 6085] <- 244950
foodprices['price'][foodprices['price'] == 29910] <- 299100
foodprices['usdprice'][foodprices['usdprice'] == 0.1449] <- 6
foodprices['usdprice'][foodprices['usdprice'] == 0.7121] <- 7
foodprices['usdprice'][foodprices['usdprice'] == 0.0641] <- 4
```

Final Data Set Overview:

Now that the data has been thoroughly cleansed, the analysis may begin.

```
summary(foodprices)
```

```
##          date                category                commodity
## Min.      :2012-04-15  cereals and tubers :154  eggs                :114
## 1st Qu.:2015-10-30  miscellaneous food :144  lentils (green):114
## Median :2020-06-15  pulses and nuts   :144  oil (vegetable):114
## Mean    :2019-02-21  meat, fish and eggs:134  rice                :114
## 3rd Qu.:2022-07-15  oil and fats      :123  sugar (white)   :114
## Max.    :2023-01-15  non-food          : 70  alcohol (spray): 10
##                               (Other)          : 90  (Other)         :279
##          unit          price          usdprice          day
## kg           :338  Min.      : 700  Min.      : 0.0166  Friday   :143
## 900 g        :144  1st Qu.: 48870  1st Qu.: 1.8333  Monday   :109
## 810 g        :123  Median : 90000  Median : 2.9253  Saturday :114
## 1.8 kg       :114  Mean    :182994  Mean    : 4.7779  Sunday   :138
## 500 g        : 30  3rd Qu.:250000  3rd Qu.: 6.1307  Thursday :138
## cubic meter: 20  Max.    :950000  Max.    :22.9927  Tuesday  :109
## (Other)     : 90                               Wednesday:108
##          month          year
## April      : 79  Min.      :2012
## August     : 79  1st Qu.:2015
## January    : 79  Median :2020
## July       : 79  Mean    :2019
## May        : 79  3rd Qu.:2022
## October    : 79  Max.    :2023
## (Other):385
```

The finalized table would be like this.

Table 2: Food Prices Table

	date	category	commodity	unit	price	usdprice	day	month	year
2	2012-04-15	cereals and tubers	rice	kg	26860	2.1909	Sunday	April	2012
3	2012-04-15	meat, fish and eggs	eggs	1.8 kg	52760	4.3034	Sunday	April	2012
4	2012-04-15	miscellaneous food	sugar (white)	kg	15890	1.2961	Sunday	April	2012
5	2012-04-15	oil and fats	oil (vegetable)	810 g	25790	2.1036	Sunday	April	2012
6	2012-04-15	pulses and nuts	lentils (green)	900 g	27369	2.2324	Sunday	April	2012

	date	category	commodity	unit	price	usdprice	day	month	year
7	2012-05-15	cereals and tubers	rice	kg	27000	2.2050	Tuesday	May	2012
8	2012-05-15	meat, fish and eggs	eggs	1.8 kg	40210	3.2838	Tuesday	May	2012
9	2012-05-15	miscellaneous food	sugar (white)	kg	15850	1.2944	Tuesday	May	2012
10	2012-05-15	oil and fats	oil (vegetable)	810 g	27880	2.2768	Tuesday	May	2012
11	2012-05-15	pulses and nuts	lentils (green)	900 g	27810	2.2711	Tuesday	May	2012

2.2 Data Analysis

As mentioned in the introduction, the current research aims to assess how US sanctions on Iran affected food prices. A trend chart is essential for our four main categories to reach this goal. A pivot table should be produced To produce this graph.

Pivot Table:

First, a new table will be created to visualize the trend change in price.

```
# New table for IRR Price group by category
foodprices_IRR <- foodprices[foodprices$date < "2022-01-01",] %>%
  group_by(date, category)%>%
  summarise(price)
```

The next step is to visualize our analysis.

```
g1 <- ggplot(data = foodprices_IRR, aes(x = date, y = price, color = category))+
  geom_line(linewidth=1)
g1+labs(title="Change Of Food Prices From 2012 To 2022 ",y="Price IRR",x="Year",
  caption="Fig.1:Food Prices Over Time") +
  scale_y_continuous(breaks=seq(10000, 700000, 100000)) +
  scale_x_date(date_labels = "%b %y")+
  geom_vline(xintercept=ymd("2018-03-08") , linetype="dotted", color = "blue") +
  annotate(geom = "text",
    x = ymd("2017-12-01") , y = 410000 ,label = "Reinstated Sanctions" ,
    angle=90 )
```

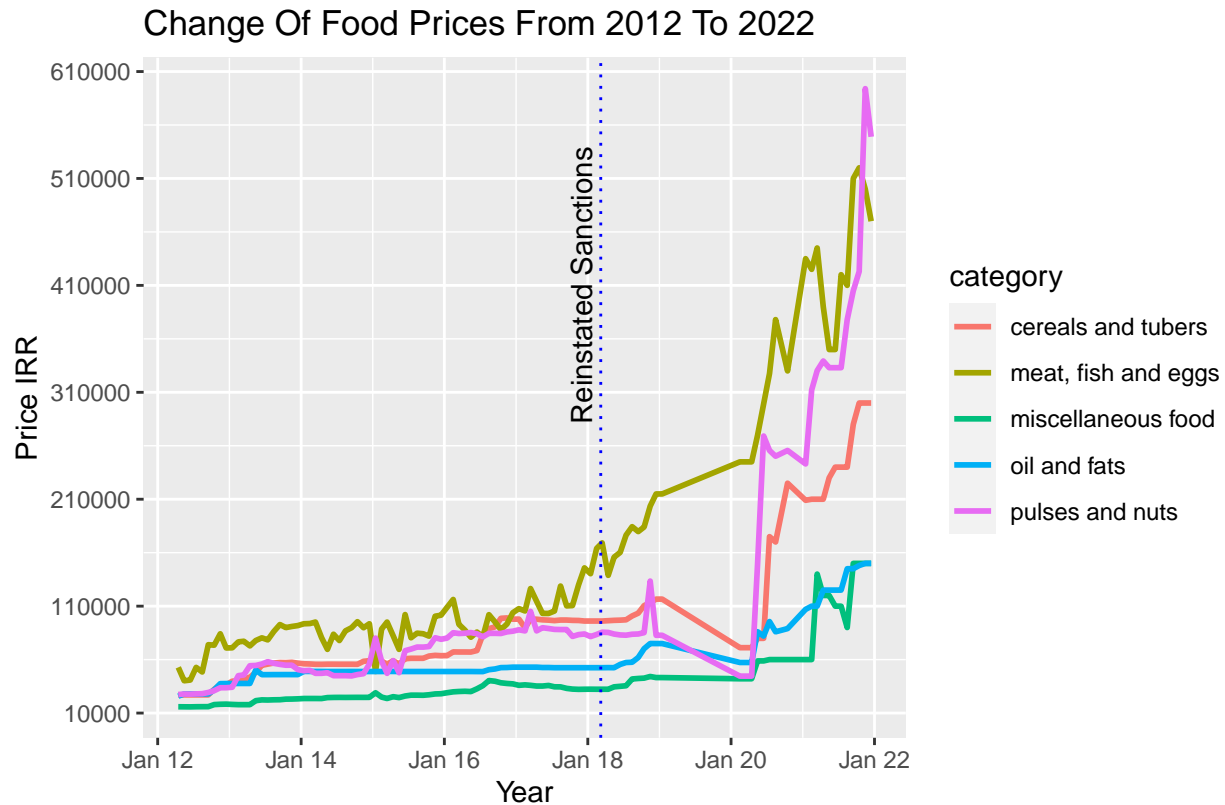


Fig.1: Food Prices Over Time

The figure depicts the price development for several food categories from 2012 to 2021. All food categories' prices have climbed significantly since the sanctions were reinstated, but the **pulses and nuts**, **meat, fish and eggs**, and **cereals and tubers** groups' price increases have been the worst.

The next step will be to demonstrate the price trend in USD currency.

```
# New table for USD Price group by category
```

```
foodprices_USD <- foodprices[foodprices$date < "2022-01-01",] %>%
  group_by(date, category)%>%
  summarise(usdprice)
```

```
g1 <- ggplot(data = foodprices_USD, aes(x = date, y = usdprice, color = category))+
  geom_line(linewidth=1)
g1 + labs(title="Change Of Food USDPrices From 2012 To 2022 ",
          y="Price USD", x="Year", caption="Fig.2: Food USDPrices Over Time") +
  scale_y_continuous(breaks=seq(1, 20, 1)) + scale_x_date(date_labels = "%b %y")+ geom_v
  annotate(geom = "text", x = ymd("2017-12-01") , y = 10 ,
          label = "Reinstated Sanctions" , angle=90 )
```


Change Of Food USDPrices From 2012 To 2022

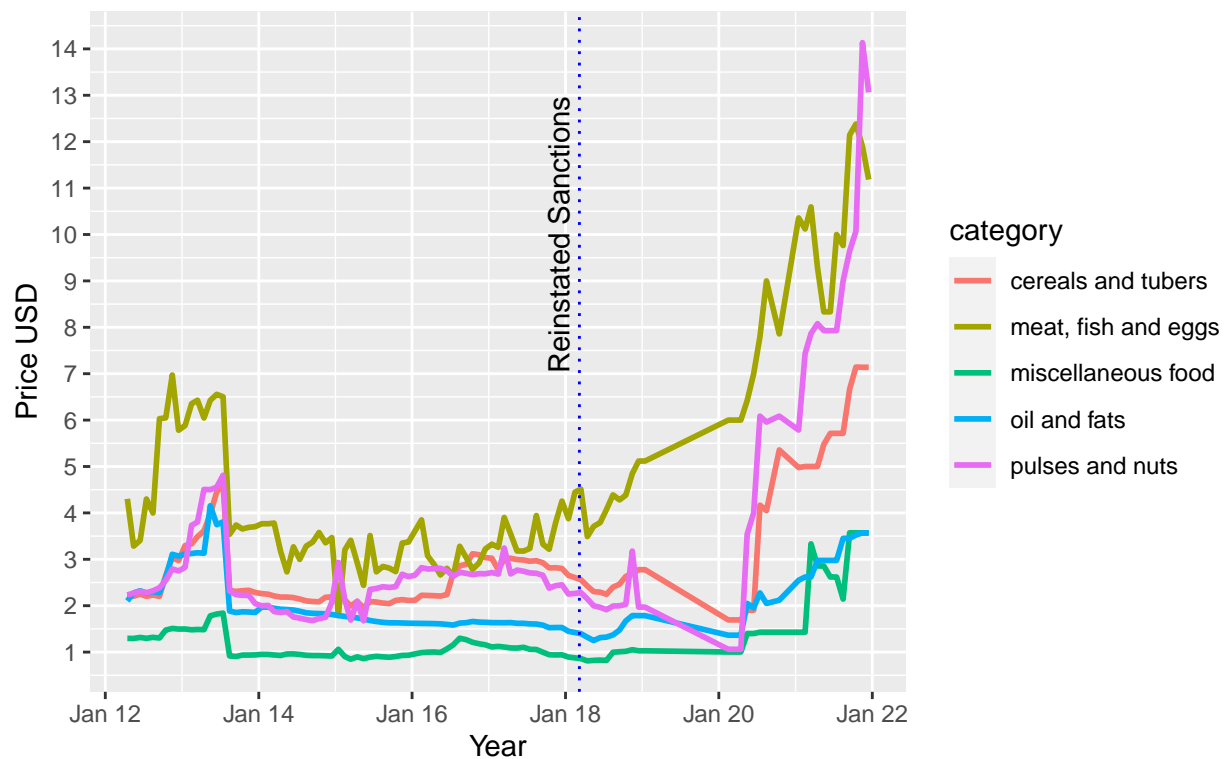


Fig.2:Food USDPrices Over Time

As was expected, the result is the same for this currency.

The third step would be to depict the exponential graph in both currencies. Because exponential functions are crucial in economics when examining growth or decay

```
foodprices_IRR$lnprice <- log(foodprices_IRR$price)
g1 <- ggplot(data = foodprices_IRR, aes(x = date, y = lnprice, color = category)) +
  geom_line(linewidth=1)
g1 + labs(title="Change Of Log Food Prices From 2012 To 2022 ",
  y="Log Price IRR", x="Year",
  caption="Fig.3:Log Food Prices Over Time") +
  scale_y_continuous(breaks=seq(8, 14, 1)) + scale_x_date(date_labels = "%b %y")+
  geom_vline(xintercept=ymd("2018-03-08") , linetype="dotted", color = "blue")+
  annotate(geom = "text", x = ymd("2017-12-01") ,
    y = 12 ,label = "Reinstated Sanctions" , angle=90 )
```

Change Of Log Food Prices From 2012 To 2022

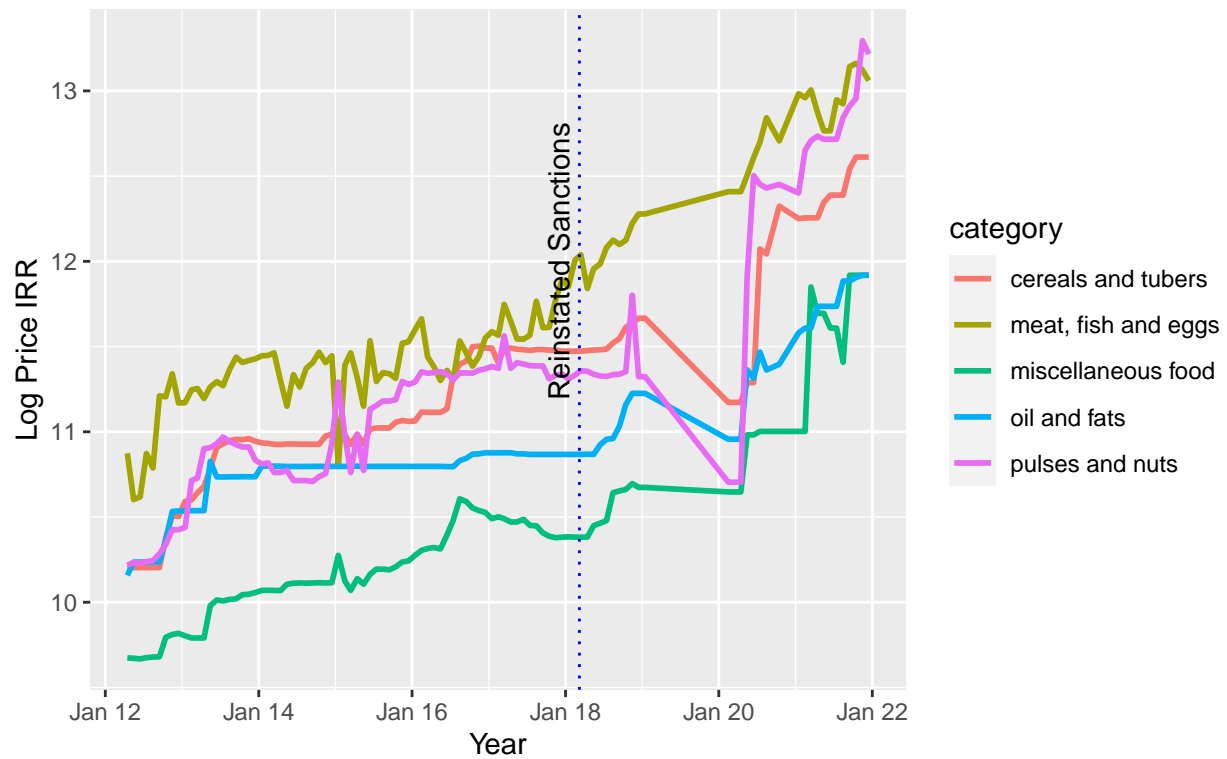


Fig.3:Log Food Prices Over Time

```
foodprices_USD$lnusdprice <- log(foodprices_USD$usdprice)
g1 <- ggplot(data = foodprices_USD,
             aes(x = date, y = lnusdprice, color = category)) +
  geom_line(linewidth=1)
g1 + labs(title="Change Of Log Food USDPrices From 2012 To 2022 ",
          y="Log Price USD", x="Year",
          caption="Fig.4:Log Food USDPrices Over Time") +
  scale_y_continuous(breaks=seq(-0.2, 2.5, 0.3)) +
  scale_x_date(date_labels = "%b %y")+
  geom_vline(xintercept=ymd("2018-03-08") , linetype="dotted", color = "blue") +
  annotate(geom = "text", x = ymd("2017-12-01") , y = 2 ,
          label = "Reinstated Sanctions" , angle=90 )
```

Change Of Log Food USDPrices From 2012 To 2022

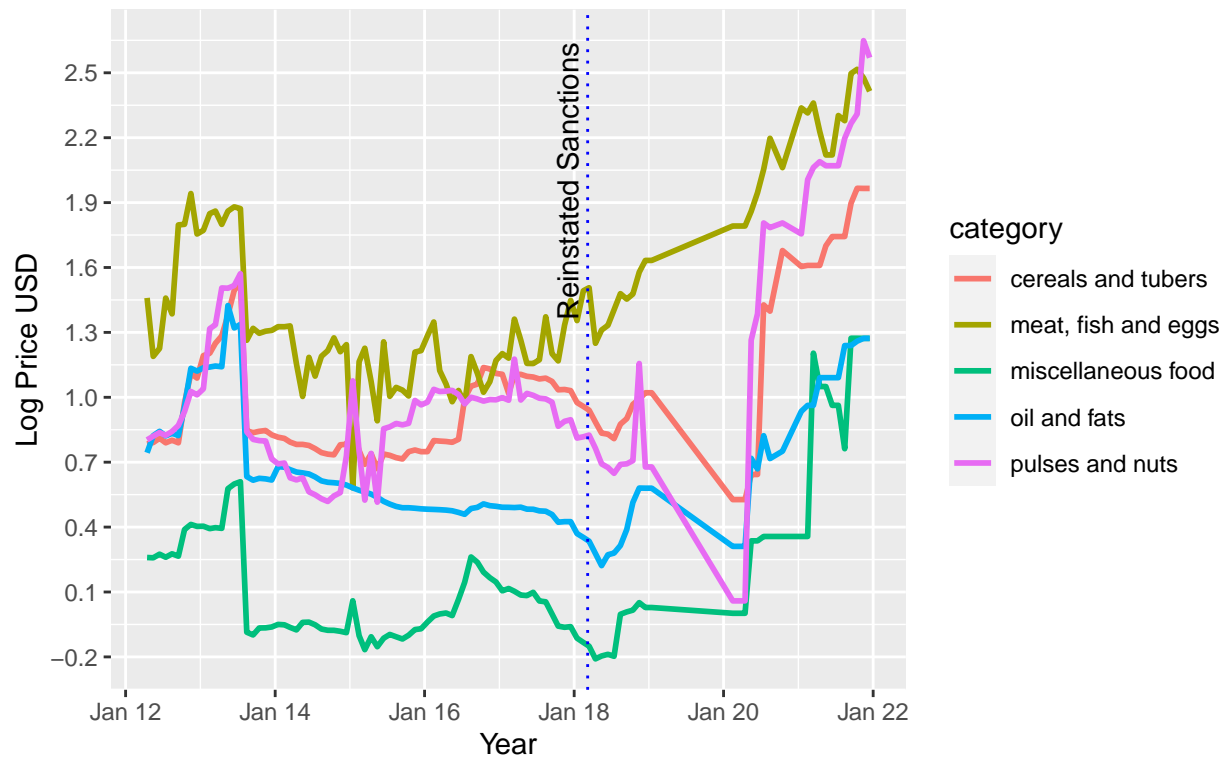


Fig.4:Log Food USDPrices Over Time

This phase involves calculating the growth rate and displaying the graph for each category. To calculate the growth rate, one should subtract the current value from the original value and divide it by the original value.

```
# New table for Log IRR Price group by category cereals and tubers
foodprices_cereals <- foodprices_IRR[foodprices_IRR$category ==
                                     "cereals and tubers",]%>%

  group_by(date, category)%>%
  summarise(lnprice)
```

```
# New table for Log IRR Price group by category meat, fish and eggs
foodprices_meat <- foodprices_IRR[foodprices_IRR$category ==
                                   "meat, fish and eggs",]%>%

  group_by(date, category)%>%
  summarise(lnprice)
```

```
# New table for Log IRR Price group by category miscellaneous food
foodprices_miscellaneous <- foodprices_IRR[foodprices_IRR$category ==
                                             "miscellaneous food",]%>%

  group_by(date, category)%>%
  summarise(lnprice)
```

```
# New table for Log IRR Price group by category oil and fats
foodprices_oil <- foodprices_IRR[foodprices_IRR$category ==
                                "oil and fats",]%>%
  group_by(date, category)%>%
  summarise(lnprice)
```

```
# New table for Log IRR Price group by category pulses and nuts
foodprices_pulses <- foodprices_IRR[foodprices_IRR$category ==
                                    "pulses and nuts",]%>%
  group_by(date, category)%>%
  summarise(lnprice)
```

The growth rate for cereals and tubers will be:

```
foodprices_cereals <- foodprices_cereals %>%
  group_by( category ) %>%
  mutate(Rate_percent = ((lnprice - lag(lnprice)) / lag(lnprice))*100)
```

```
ggplot(foodprices_cereals) +
  geom_line(linewidth=1) +
  aes(x = date , y = Rate_percent , color = category) +
  geom_line() +
  scale_y_continuous(breaks=seq(-7, 7, 0.5)) +
  labs(title="Growth rate for cereals From 2012 To 2022 ",
       caption="Fig.5:Growth rate for cereals")
```

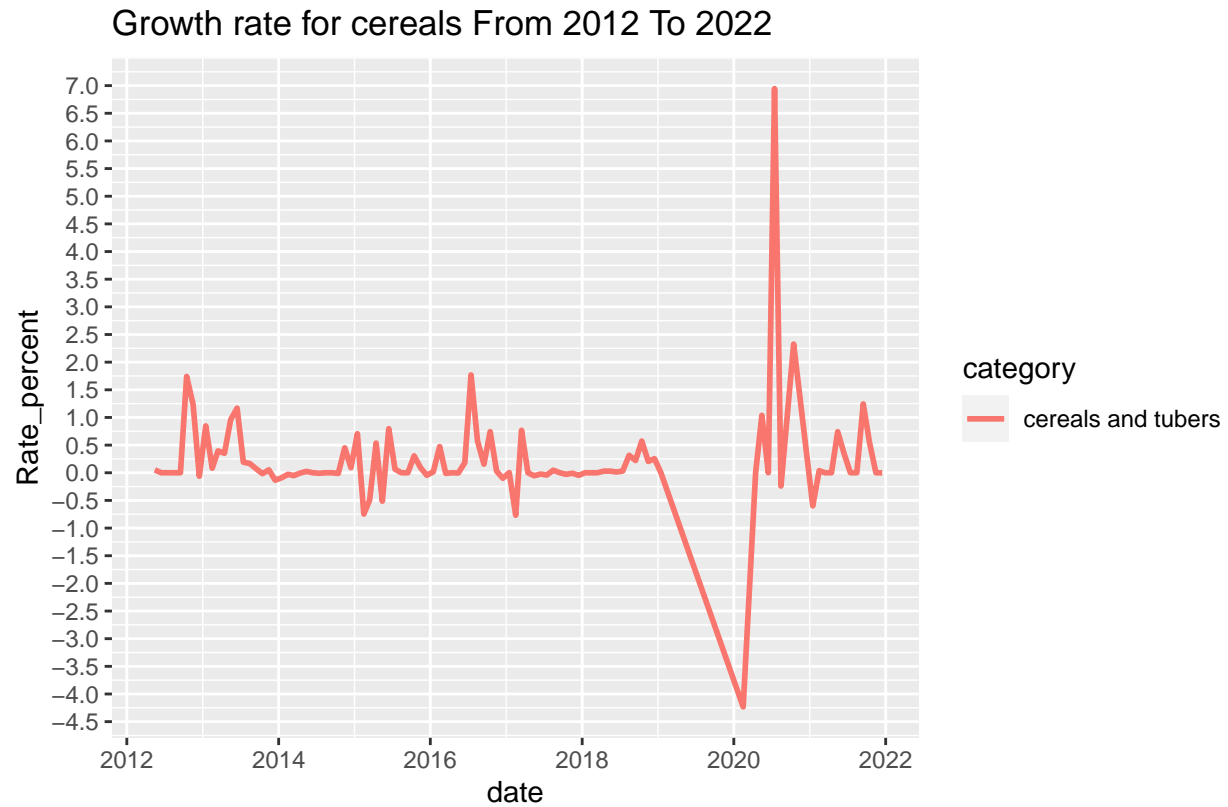


Fig.5:Growth rate for cereals

The growth rate for meat, fish and eggs will be:

```
foodprices_meat <- foodprices_meat %>%
  group_by( category ) %>%
  mutate(Rate_percent = ((lnprice - lag(lnprice)) / lag(lnprice))*100)
```

```
ggplot(foodprices_meat) +
  geom_line(linewidth=1) +
  aes(x = date , y = Rate_percent , color = category) +
  geom_line() +
  scale_y_continuous(breaks=seq(-7, 7, 0.5)) +
  labs(title="Growth rate for meat From 2012 To 2022 ",
       caption="Fig.6:Growth rate for Meat")
```

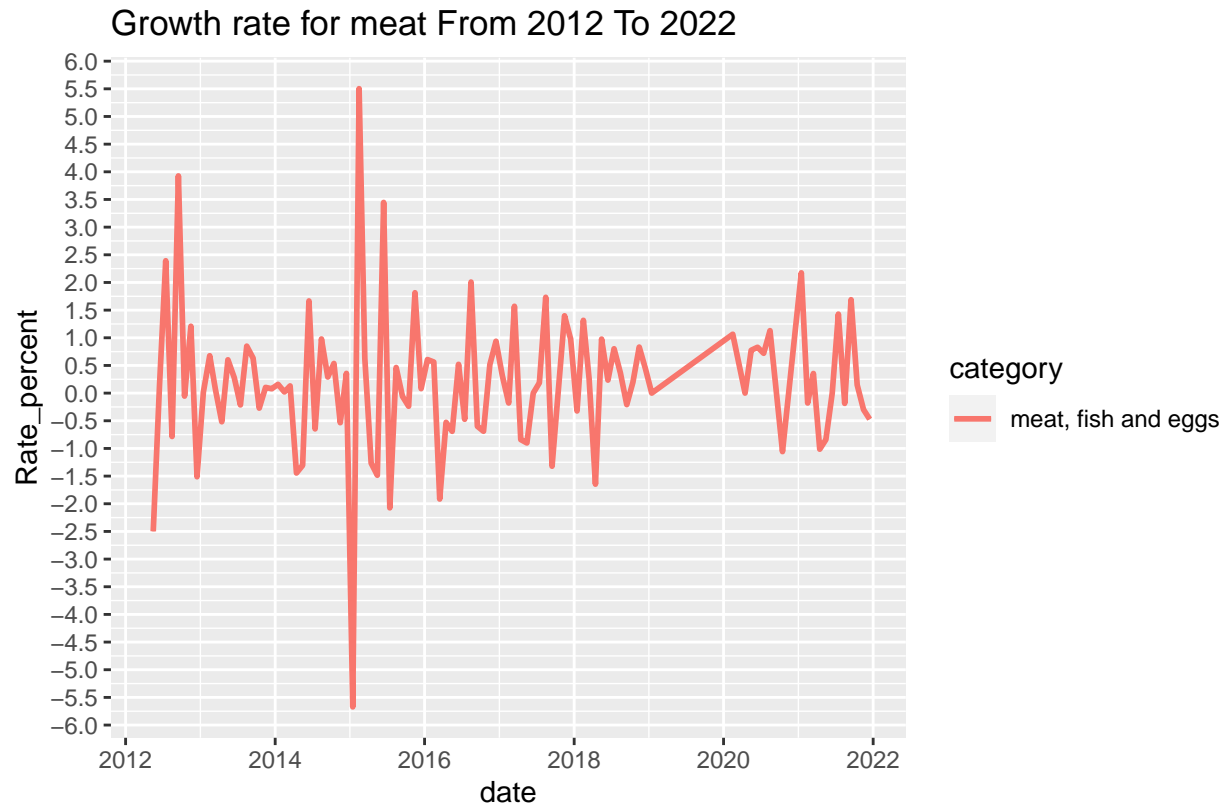


Fig.6:Growth rate for Meat

The growth rate for miscellaneous food will be:

```
foodprices_miscellaneous <- foodprices_miscellaneous %>%
  group_by( category ) %>%
  mutate(Rate_percent = ((lnprice - lag(lnprice)) / lag(lnprice))*100)
```

```
ggplot(foodprices_miscellaneous) +
  geom_line(linewidth=1) +
  aes(x = date , y = Rate_percent , color = category) +
  geom_line() +
  scale_y_continuous(breaks=seq(-7, 8, 0.5)) +
  labs(title="Growth rate for miscellaneous From 2012 To 2022 ",
       caption="Fig.7:Growth rate for miscellaneous")
```

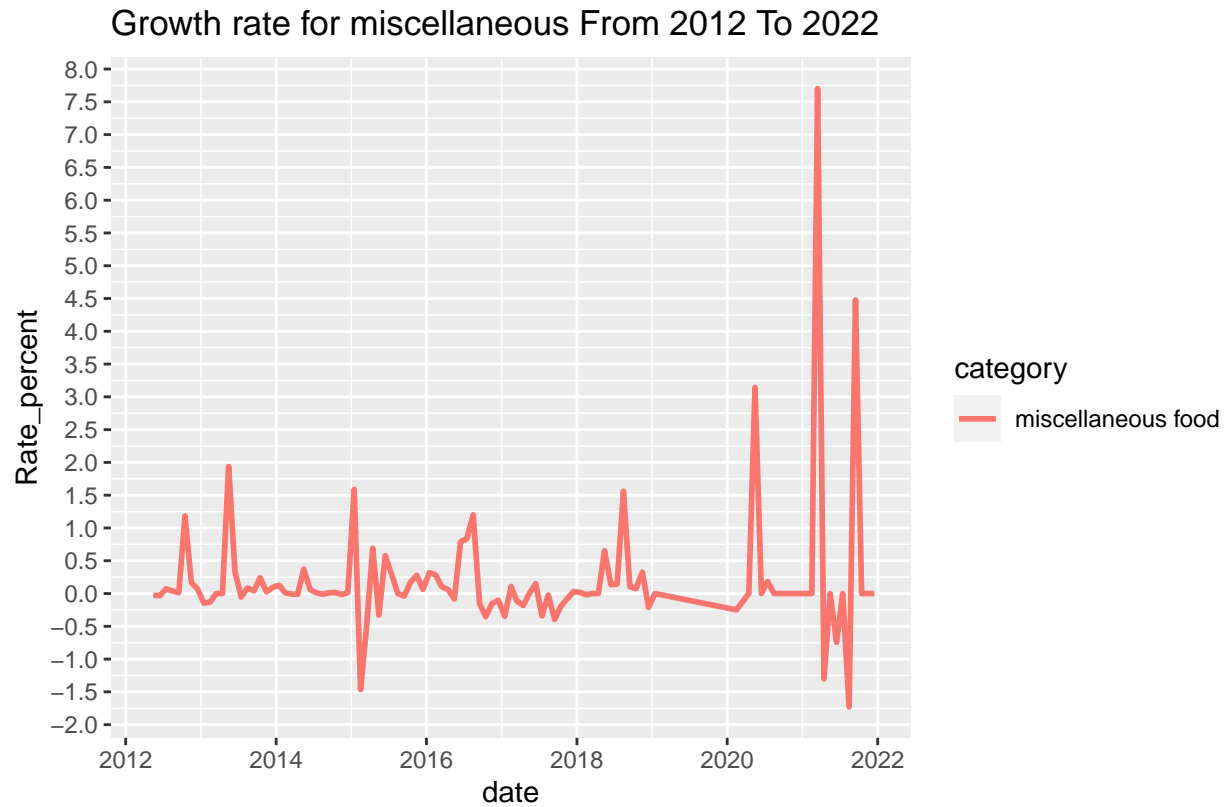


Fig.7:Growth rate for miscellaneous

The growth rate for oil and fats food will be:

```
foodprices_oil <- foodprices_oil %>%
  group_by( category ) %>%
  mutate(Rate_percent = ((lnprice - lag(lnprice)) / lag(lnprice))*100)
```

```
ggplot(foodprices_oil) +
  geom_line(linewidth=1) +
  aes(x = date , y = Rate_percent , color = category) +
  geom_line() +
  scale_y_continuous(breaks=seq(-7, 8, 0.5)) +
  labs(title="Growth rate for oil and fat From 2012 To 2022 ",
       caption="Fig.8:Growth rate for oil and fat")
```

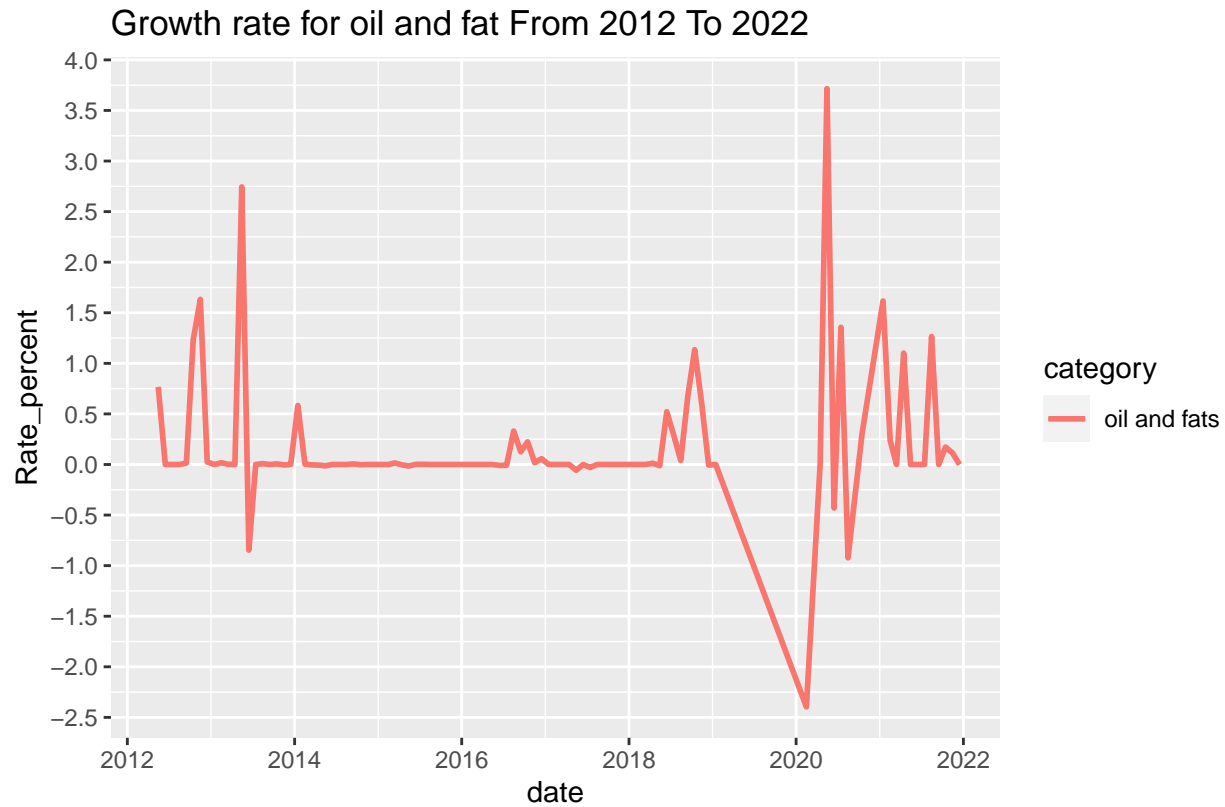


Fig.8:Growth rate for oil and fat

The growth rate for pulses and nuts food will be:

```
foodprices_pulses <- foodprices_pulses %>%
  group_by( category ) %>%
  mutate(Rate_percent = ((lnprice - lag(lnprice)) / lag(lnprice))*100)
```

```
ggplot(foodprices_pulses) +
  geom_line(linewidth=1) +
  aes(x = date , y = Rate_percent , color = category) +
  geom_line() +
  scale_y_continuous(breaks=seq(-7, 12, 1)) +
  labs(title="Growth rate for pulses and nuts From 2012 To 2022 ",
       caption="Fig.9:Growth rate for pulses and nuts")
```

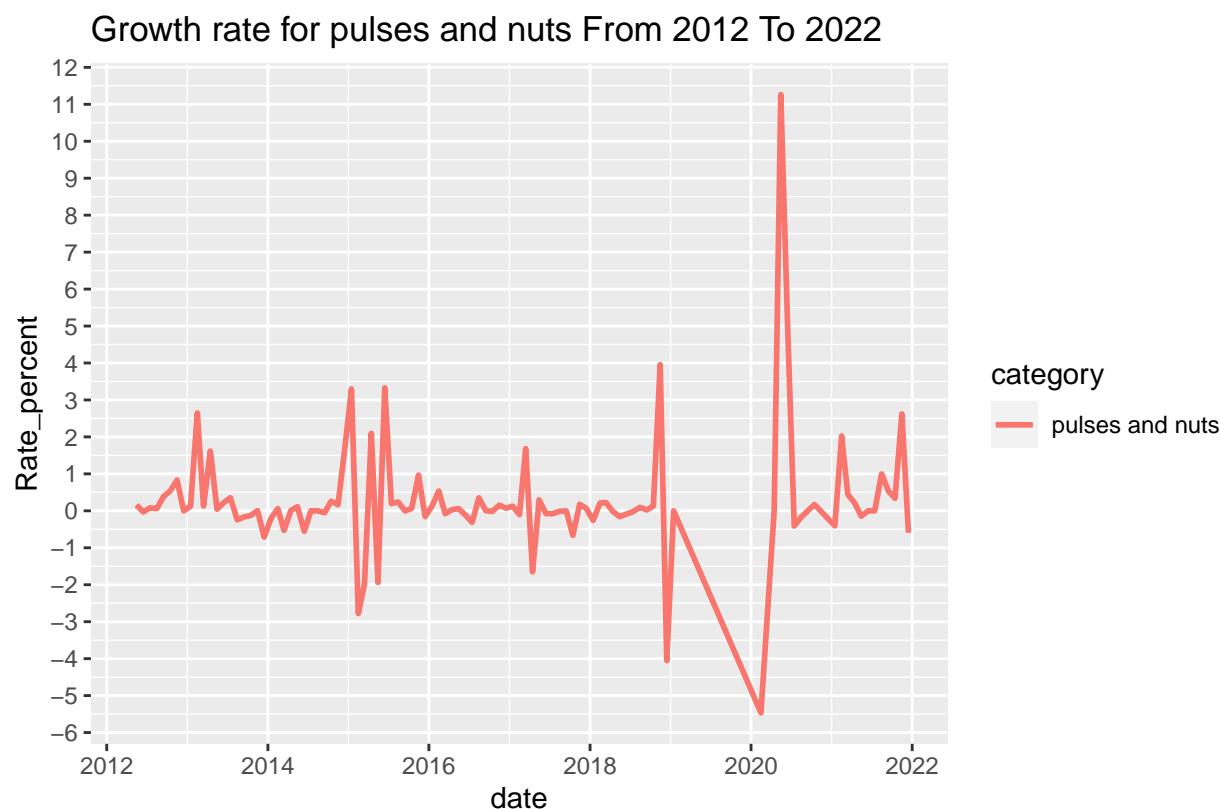



Fig.9:Growth rate for pulses and nuts

3 Discussion

Our data shows that over the years after the sanctions, the prices of the majority of food goods climbed by more than 300%. Those that are rich in nutrition, like nuts and meat products, have the greatest inflation rates, whereas foods that are poor in nutrition, such as oil and fat, have the lowest inflation rates.

According to a comprehensive study by (Kokabisaghi 2018), shipping products to Iran, including food and medicines, had become very difficult and costly as a result of the sanctions, on trade, banking, the financial system, and shipping. According to a research, these restrictions have led to a 40% rise in the cost of trade (CJ.Gordon 2012). In 2012 and following the first round of nuclear sanctions against Iran, the United Nations Children's Fund (UNICEF) described Iran as being subject to strict forcible economic sanctions that had a harmful impact on the environment, healthcare, and the social economic factors that influence the health of common people, especially kids (UNICEF 2012).

It is commonly known that nutrient-rich meals cost more than those high in empty calories, and therefore a healthy diet is more expensive than an unhealthy one (N. Darmon 2015). According to a research, the two food categories that make up a healthy diet—meats and fruits—are the most expensive, while the food category that costs the least is the fats (J. Hejazi 2022). Due to the present economic climate and government subsidies for particular

goods, including white bread, Iranians will be compelled to eat energy-dense meals like fats, sweets, and refined grains. As a consequence of consuming this unhealthful diet, an unprecedented rise in obesity and its associated ailments, such as cancer, diabetes, and cardiovascular disease, is unavoidable soon (DP. Reidlinger 2017). Research on the impacts of sanctions imposed by the United Nations (UN) or the US on the food security of inhabitants of sanctioned nations is scarce; nonetheless, studies from Iraq (Field 1993) and Cuba (Ross 2004) may attest to the detrimental effects of sanctions on those countries' populations' access to food.

Following the reinstatement of US nuclear restrictions, the Iranian government implemented specific measures to lessen the harsh impact of the penalties on the general populace. In April 2018, the Central Bank of Iran granted Iranian businesses importing certain “critical commodities” a discounted rate of 42 000 Iranian rials per US dollar (the dollar now fetches more than 480 000 rials on Iran’s open market). Some food products, such as different kinds of meats, fish, egg, oils, rice, etc., were classified as necessities; however, it may be inferred from the pricing of these goods in the years that followed the reimposition of sanctions that the Iran Central Bank’s effort had little to no impact on regulating food prices, if any. This claim’s veracity is supported by the fact that meat prices roughly tripled on average throughout these times. Thus, the Iranian government needs to evaluate the effectiveness of this approach.

The Iranian government can take steps to lessen the harmful effects of the sanctions on Iranian households by developing effective food assistance programs, establishing food banks with the help of charities and non-governmental organizations, and then implementing nutrition education programs to teach people how to make the right food choices by their budget (C. Bazerghi 2016).

It is pretty evident that, as with many other ongoing wars in the globe, regular Iranian citizens—particularly those who belong to at-risk demographics like children and the elderly—are the only ones to suffer as a result of the present political tensions between the United States and Iran.

4 Conclusion

Due to the enormous food market inflation that followed the re-imposition of US sanctions on Iran and the low income of Iranians, most of the population needs help to maintain an average healthy diet. This might make Iranians more susceptible to chronic illnesses soon, and if this pattern continues, it could put the nation at risk of a food crisis.

5 Limitations and Challenges

This study has some limitations. First, the database includes just five products, which is insufficient for precise analysis and could not bring comprehensive results. Secondly, this study does not consider the inflation rate and exchange rate of IRR against the US dollar.

It would be better to compare the general national inflation rate with inflation in the food sector and the growth rate of each product with changes in the exchange rate. Furthermore, it could have more analysis of growth rate diagrams. Third, for further study, consider the government's financial policies in these years and investigate if there were better solutions to diminish the effect of sanctions on important sections like food.

However, I confronted many challenges in implementing this analysis, from installing Rstudio to executing the final file. I decided to mention some most time-consuming challenges that I had. One of them was the shortcode for calculating the words in the project. It did not execute; after many searches, it was solved just by an update. I needed to understand if all products in the time series have the same unit, so I asked questions in Stackoverflow, and somebody helped me with this problem. Also, illustrating diagrams was a challenge and took more time than I expected. Finally, calculating the growth rate was a challenge that I spent time on it.

Also, all related files of this study were uploaded to GitHub

References

- Bloomberg. 2022. "Russia Is Now the World's Most-Sanctioned Nation." *Bloomberg.com*.
- Byrne, Ciara. 2017. *Development Workflows for Data Scientists*. O'Reilly Media.
- C. Bazerghi, M. Dunn, FH. McKay. 2016. "The Role of Food Banks in Addressing Food Insecurity: A Systematic Review." *J Community Health* 41 (4): 732–40. <https://doi.org/10.1007/s10900-015-0147-5>.
- CJ.Gordon. 2012. "Crippling Iran: The UN Security Council and the Tactic Ofdeliberate Ambiguity." *Georget J Int Law* 44 (3): 973–1006.
- Collier, Ken. 2012. *Agile Analytics: A Value-Driven Approach to Business Intelligence and Data Warehousing*. Addison-Wesley.
- Conway, D. 2023. "The Data Science Venn Diagram." <http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>.
- Crowston, Kevin, Jeff S Saltz, Amira Rezgui, Yatish Hegde, and Sangseok You. 2019. "Socio-Technical Affordances for Stigmergic Coordination Implemented in MIDST, a Tool for Data-Science Teams." *Proceedings of the ACM on Human-Computer Interaction* 3 (CSCW): 1–25.
- Das, Manirupa, Renhao Cui, David R. Campbell, Gagan Agrawal, and Rajiv Ramnath. 2015. "Towards Methods for Systematic Research on Big Data." In *2015 IEEE International Conference on Big Data (Big Data)*, 2072–81. <https://doi.org/10.1109/BigData.2015.7363989>.
- Database, World Food Programme Price. 2022. "Iran (Islamic Republic of) - Food Prices." *Humdata*. <https://data.humdata.org/dataset/wfp-food-prices-for-iran-islamic-republic-of>.
- DP. Reidlinger, LM. Goff, TA. Sanders. 2017. "How Expensive Is a Cardioprotective Diet? Analysis from the CRESSIDA Study." *Public Health Nutr.* 20 (8): 1423–30. <https://doi.org/10.1017/S1368980016003529>.
- ecfr.gov. 2010. "31 CFR § 560540 - Exportation of Certain Services and Software Incident

- to Internet-Based Communications.” *Code of Federal Regulations (Ecf)*. <https://www.ecfr.gov/current/title-31/subtitle-B/chapter-V/part-560/subpart-E/section-560.540>.
- Field, JO. 1993. “From Food Security to Food Insecurity: The Case of Iraq 1990-91.” *Geo Journal* 30 (2): 185–94. <https://doi.org/10.1007/BF00808137>.
- Grady, Nancy W. 2016. “KDD Meets Big Data.” In *2016 IEEE International Conference on Big Data (Big Data)*, 1603–8. <https://doi.org/10.1109/BigData.2016.7840770>.
- Granville, V. 2014. *Developing Analytic Talent: Becoming a DataScientis*. 1st ed. Wiley Publishing.
- humdata. 2022. “Data Licenses.” *Humdata*. <https://data.humdata.org/faqs/licenses>.
- J. Hejazi, S. Emamgholipour. 2022. “The Effects of the Re-Imposition of US Sanctions on Food Security in Iran.” *Int J Health Policy Manag* 11 (5): 651–57. [10.34172/ijhpm.2020.207](https://doi.org/10.34172/ijhpm.2020.207).
- JCPOA. 2019. “EU Officially Announces October 18 Adoption Day of JCPOA.” *Islamic Republic News Agency*. <https://en.irna.ir/news/81804426/EU-officially-announces-October-18-adoption-day-of-JCPOA>.
- Josh Levs, CNN. 2012. “A Summary of Sanctions Against Iran.” Online. <https://edition.cnn.com/2012/01/23/world/meast/iran-sanctions-facts/index.html>.
- Jurney, Russell. 2017. *Agile Data Science 2.0: Building Full-Stack Data Analytics Applications with Spark*. OReilly Media, Inc.
- Kaufmann, Michael. 2019. “Big Data Management Canvas: A Reference Model for Value Creation from Data.” *Big Data and Cognitive Computing* 3 (1). <https://doi.org/10.3390/bdcc3010019>.
- Kokabisaghi, F. 2018. “Assessment of the Effects of Economic Sanctions on Iranians Right to Health by Using Human Rights Impact Assessment Tool.” *Int J Health Policy Manag* 7 (5): 374–93. http://www.ijhpm.com/article_3454.html.
- Lab, Domino Data. 2017. “Managing Data Science Teams.”
- Lakshmanan, Indira A. R. 2013. “U.s. Senators Seeking Tougher Economic Sanctions on Iran.” *Bloomberg News*. [https://www.bloomberg.com/news/articles/2013-04-09/u-s-senators-seeking-tougher-economic-sanctions-on-iran?leadSource=uverify wall](https://www.bloomberg.com/news/articles/2013-04-09/u-s-senators-seeking-tougher-economic-sanctions-on-iran?leadSource=uverify%20wall).
- Larson, Deanne, and Victor Chang. 2016. “A Review and Future Direction of Agile, Business Intelligence, Analytics and Data Science.” *International Journal of Information Management* 36 (5): 700–710.
- Loukide, M. 2011. *What Is Data Science?* OReilly Media, Inc.
- Martinez, Inigo, Elisabeth Viles, and Igor Olaizola. 2021. “Data Science Methodologies: Current Challenges and Future Approaches.” *Big Data Research* 24 (January): 100183. <https://doi.org/10.1016/j.bdr.2020.100183>.
- Moyle, Steve, and Alípio Jorge. 2001. “Ramsys-a Methodology for Supporting Rapid Remote Collaborative Data Mining Projects.” In *ECML/PKDD01 Workshop: Integrating Aspects of Data Mining, Decision Support and Meta-Learning (IDDM-2001)*. Vol. 64.
- N. Darmon, A. Drewnowski. 2015. “Contribution of Food Prices and Diet Cost to Socio Economic Disparities in Diet Quality and Health: A Systematic Review and Analysis.” *Nutr Rev* 73 (10): 643–60. <https://doi.org/10.1093/nutrit/nuv027>.
- Nichols, Louis, Michelle & Charbonneau. 2012. “U.n. Chief Says Sanctions on Iran Affecting Its People.” *Reuters*. <https://www.reuters.com/article/us-iran-sanctions-un-idUSBRE89412Z20121005>.

- Ross, JE. 2004. "Food Security in Cuba." *Cuba Today:Continuity and Change Since the Periodo Especial.*, 115–25.
- Saltz, Jeff, Nicholas Hotz, David Wild, and Kyle Stirling. 2018. "Exploring Project Management Methodologies Used Within Data Science Teams."
- Saltz, Jeffrey, Frank Armour, and Ramesh Sharda. 2018. "Data Science Roles and the Types of Data Science Programs." *Communications of the Association for Information Systems* 43 (1): 33.
- Shearer, Colin. 2000. "The CRISP-DM Model: The New Blueprint for Data Mining." *Journal of Data Warehousing* 5 (4): 13–22.
- UNICEF. 2012. "Annual Report 2012 for Iran (MENA)." *UNICEF*.
- Vanauer, Martin, Carsten Bohle, and Bernd Hellingrath. 2015. "Guiding the Introduction of Big Data in Organizations: A Methodology with Business- and Data-Driven Ideation and Enterprise Architecture Management-Based Implementation." In *2015 48th Hawaii International Conference on System Sciences*, 908–17. <https://doi.org/10.1109/HICSS.2015.113>.
- Warden, P. 2011. "Why the Term Data Science Is Flawed but Useful." *OReilly Radar*.
- Wikipedia. 2022. "Sanctions Against Iran." *Wikipedia*. https://en.wikipedia.org/wiki/Sanctions_against_Iran. https://en.wikipedia.org/wiki/Sanctions_against_Iran#Iran_was_the_most_sanctioned,in_Tehran_and_took_hostages.
- Wills, Josh. 2012. "Data Scientist Definition."
- Zirulnick, Ariel. 2011. "Sanction Qaddafi? How 5 Nations Have Reacted to Sanctions." *The Christian Science Monitor*. <https://www.csmonitor.com/World/Global-Issues/2011/0224/Sanction-Qaddafi-How-5-nations-have-reacted-to-sanctions./Iran>.