Forecasting Monthly Vegetable Prices

Jaime Neil P. Alap

Glomie G. Gonzales

Eloisa J. Jimenez

Carlo D. Pastores, Jr.

# Chapter I

## INTRODUCTION

### Background of the Study

The fluctuation of vegetable prices is a global concern in both developed and emerging economies. The need for vegetables as a primary source of essential nutrients and dietary variety has grown as urbanization and global population growth simultaneously. Consequently, the dynamics of vegetable pricing have grown more intricate and significant, having an effect on not just consumer choices but also food security, economic stability, and agricultural practices. Understanding the factors driving these price fluctuations is paramount for policymakers, farmers, consumers, and the broader food industry (Chen et al., 2018).

Vegetables are a major life necessity for urban and rural residents, and the vegetable market massively supports rural economic development. Fluctuations in vegetable prices affect farmers’ income, quality of life, and decision-making regarding vegetable planting. Therefore, issues regarding maintaining price stability have long been focused on government policies. In recent years, the price of vegetables has exhibited dramatic and frequent volatility, which caused a series of negative effects on stakeholders in the supply-chain, e.g., farmers, logistics, wholesale, retail, and consumers. In view of the above adverse effects on stakeholders, it is of crucial importance to filter out key factors that relate to price fluctuations, targeting effective monitoring of real-time abnormal fluctuations.

Price fluctuation is a complex issue with far-reaching consequences, especially for vulnerable populations. While higher prices may appear beneficial for farmers, the inherent volatility poses significant risks, potentially leading to substantial losses for agricultural stakeholders. Often attributed to imbalances in market fundamentals, where demand surpasses supply, this phenomenon has a profound impact on small-scale farmers (Mchopa et al., 2014).

According to *FAOSTAT* (n.d.), agricultural products account for a large proportion of the market as a necessity for daily consumption, and their prices play a critical part in consumer spending and agricultural household income. The supply and demand in a given year determine the prices of agricultural products. While an undersupply of agricultural items raises prices and burdens consumers, an oversupply of agricultural products causes vegetable prices to increase and causes financial losses to farming households.

In the Philippines, vegetables price height has been a major problem for past years. Consumers were complaining of higher food prices, reflected not only in vegetables but in poultry and pork meat as well, and selected fish variants. Some prices have gone up by as much as 66 percent from last month alone. The Department of Agriculture officials have blamed the higher prices on numerous factors such as the ongoing pandemic, the devastating typhoons that destroyed crops, and the unwillingness of poultry and hog raisers to farm a new following the glut and the spread of the African swine fever, respectively. (Ocampo, 2021)

As stated by Vibas and Raqueño (2019), Agricultural commodities significantly impact a country’s export earnings and economic performance. Price fluctuations affect farmers, consumers, and public agencies. The Department of Agriculture in the Philippines recognizes that local market-driven commodity pricing, notably for fruits and vegetables, have prompted government action.

Nueva Vizcaya Agricultural terminal Inc. general manager Gilbert Cumila said, Vegetable prices in the Philippines continue to rise as demand exceeds supply. Wholesale prices of vegetables increases due to the demand of people. This significantly affect the way consumers purchase vegetables to sustain their necessities.

Vegetable prices in Nueva Vizcaya is rapidly increasing month by month. Consumers are having hard time to provide necessities due to this reason. This study will forecast the changes of vegetable prices in Nueva Vizcaya to serve as guide to the consumers directly affected by the problem.

This research aims to describe the monthly vegetable prices in Nueva Vizcaya and identify the models that will forecast monthly vegetables. The research oath to determine the best model that can be used to forecast monthly vegetable prices in Nueva Vizcaya.

### Statement of the Problem

The purpose of this study was to examine ten-year historical data on monthly vegetable prices in Nueva Vizcaya on the viability of NVAT, Bambang, Nueva Vizcaya. To achieve this, the researchers will first provide a comprehensive description of the historical trends and fluctuations in monthly vegetable prices in the region. Subsequently, it will explore various forecasting models to identify the most suitable ones for predicting these prices effectively. The primary objective is to determine the best-performing model among the options considered. Once the optimal model is identified, it will be used to generate forecasts for monthly vegetable prices in Nueva Vizcaya, thereby assisting local farmers, distributors, and consumers in making informed decisions related to agricultural production, distribution, and purchasing of vegetables in the region. This research addresses a critical issue of food security and economic stability in Nueva Vizcaya and can have broader implications for agricultural forecasting in similar regions.

### Objectives of the Study

The researchers will forecast monthly vegetable prices in Nueva Vizcaya by accomplishing the following:

1. Describe the monthly vegetable prices in Nueva Vizcaya.
2. Identify models that will forecast monthly vegetable prices in Nueva Vizcaya.
3. Determine the best model that will forecast monthly vegetable prices in Nueva Vizcaya.
4. Forecast the best model for monthly vegetable prices in Nueva Vizcaya.

### Significance of the Study

This study focused on determining the best model that will forecast monthly vegetable prices in Nueva Vizcaya on the viability of NVAT, Bambang, Nueva Vizcaya.

**Department of Agriculture**. This research would be beneficial to the Department of Agriculture by incorporating its findings into policy formulation and resource allocation. Accurate price forecasts can inform agricultural development programs, subsidies, and interventions, ultimately supporting the growth and sustainability of the agricultural sector in Nueva Vizcaya.

**Farmers**. This research is beneficial to farmers as it holds immense significance for farmers in Nueva Vizcaya and provides them with valuable insights into future vegetable price trends. Accurate price forecasts enable farmers to plan their planting and harvesting schedules efficiently, reduce wastage, and optimize their crop yields, ultimately leading to improved income stability and sustainable agricultural practices.

**Vendors**. This research would be beneficial to vendors in Nueva Vizcaya’s vegetable markets by obtaining reliable price forecasts. With this information, they can make informed purchasing decisions, maintain competitive prices, and increase profit margins. This, in turn, fosters a more stable and profitable business environment for vendors.

**Consumers**. This research would be beneficial to consumers as it helps maintain price stability and affordability. When vendors can make better decisions based on accurate forecasts, consumers are less likely to experience price shocks or sudden increases in vegetable prices, ensuring accessibility to essential food items.

**Future Entrepreneurs**. This research would be beneficial to future entrepreneurs as they want to enter the vegetable market in Nueva Vizcaya, and they can use the results of this study to make informed business decisions. This information helps develop market entry strategies, inventory management strategies, and pricing strategies to reduce the risks associated with launching a new business. Business Owners. This research would be beneficial to business owners in Nueva Vizcaya, whether they are in the agricultural sector or the retail and distribution sector, and they can leverage predictive insights to optimize their supply chain operations. Improved supply chain efficiency can lead to cost savings and increased profitability.

**Future Researchers**. This research serves as a valuable foundation for future researchers interested in agricultural economics, market dynamics, and forecasting methodologies. It provides a benchmark dataset and insights that can be expanded upon and refined in subsequent research efforts, contributing to the continuous advancement of agricultural forecasting and market analysis.

### Scope and Delimitation of the Study

The study focuses on forecasting the monthly prices of vegetables in the NVAT, Bambang, Nueva Vizcaya, for the years 2013 to 2023. The researchers decided that the Box and Jenkins models were the best strategy for forecasting monthly prices of vegetables in NVAT, Bambang, Nueva Vizcaya. However, it is vital to highlight that the study’s findings and conclusions are limited to this specific locale, the vegetables, and the timeframe mentioned. Alternative forecasting methodologies and external variables impacting vegetable prices are not investigated in this study. As a result, the findings should be regarded with caution and applied exclusively within the defined boundaries of NVAT, Bambang, Nueva Vizcaya, and the chosen vegetables during the specified study period.

### Conceptual Framework

The figure….

### Definition of Terms

**Akaike Information Criterion (AIC)**. AIC is a statistical measure used to compare the goodness of fit of different statistical models. It balances the complexity of a model with its ability to explain the data. Lower AIC values suggest better model fit.

**ARIMA (AutoRegressive Integrated Moving Average)**. ARIMA is a popular time series forecasting model that combines autoregressive (AR) and moving average (MA) components with differencing to make a time series stationary. It is used for modeling and forecasting univariate time series data.

**Best Fit Model**. The best fit model is the statistical or mathematical model that most accurately describes and predicts the patterns and trends within a given time series. It is selected based on criteria like RMSE, AIC, BIC, or other statistical measures.

**Box-Jenkins Model**. The Box-Jenkins methodology, also known as the ARIMA modeling process, is a systematic approach for identifying, estimating, and forecasting time series data using autoregressive integrated moving average models. It involves model identification, estimation, diagnostic checking, and forecasting.

**Exponential Smoothing**. Is a time series forecasting method for univariate data that can be extended to support data with a systematic trend or seasonal component. It is a powerful forecasting method that may be used as an alternative to the popular Box-Jenkins ARIMA family of methods.

**Forecasting**. Is the process of making predictions or estimates about future values or events based on historical data and patterns. It is widely used in various fields, including economics, finance, and weather prediction.

**Inflation**. It refers to a broad rise in the prices of goods and services across the economy over time, eroding purchasing power for both consumers and businesses.

**The Nueva Vizcaya Agricultural Terminal (NVAT), Philippines**. NVAT is a mixed capital public-private joint venture established in Nueva Vizcaya in 2004 to address problems such as lack of markets and keen market competition.

**Root Mean Squared Error (RMSE)**. RMSE is a measure of the average deviation between the values predicted by a model or forecast and the actual observed values in a time series. It quantifies the accuracy of a predictive model, with lower values indicating better predictive performance.

**SARIMA (Seasonal AutoRegressive Integrated Moving Average)**. SARIMA is an extension of the ARIMA model that includes seasonal components to account for seasonality in time series data. It is particularly useful for data with recurring seasonal patterns.

**Time Series**. A time series is a sequence of data points collected or recorded at specific time intervals. It is used to analyze and predict trends, patterns, and behavior over time.

**Vegetable Price**. Vegetable price refers to the cost or price at which vegetables are bought or sold in a market or region. It is a key indicator in economics and agriculture, reflecting supply and demand dynamics.

# Chapter II

## REVIEW OF RELATED LITERATURE

**Vegetable**

The term “vegetable” typically refers to the living, edible components of some herbaceous plants, such as their roots, stems, leaves, flowers, fruit, or seeds. These plant components are eaten raw or cooked in a number of ways, primarily as a savory dish rather than a dessert. Vegetables are low in calories, include more than 70% water, only have a little amount of protein (about 3.5%), and less than 1% fat. These are a good source of minerals like calcium and iron as well as vitamins like A and C. Almost all vegetables are rich in dietary fiber and antioxidants (Vegetable | Description, Types, Farming, & Examples, 2023).

Vegetables have various types, and these include cabbage, stalk vegetables, leafy vegetables, salad greens, vegetable fruits, bulbs, mushrooms, tubers, and specialty vegetables. Stalk vegetables are made from cellulose-rich plant stalks, whereas cabbage vegetables are those grown for their heads, leaves, or flowers. Among the salad greens grown for their edible leaves are endives, lettuce stems, edible pods, and immature shoots. Additionally, specialty vegetables, tubers, and fungi are included in this categorization (Rajak, 2023).

Vegetables are full of essential vitamins, minerals, and antioxidants that provide many important health benefits to your body. For instance, carrots are known for being very high in vitamin A, which plays an important role in eye health, as you grow older.

Vegetables also offer many other health benefits like:

**1. Improved Digestive Health**

Vegetables are a good source of dietary fiber, a type of carbohydrate that helps pass food through your digestive system. Studies show that fiber may also improve vitamin and mineral absorption in the body, which could potentially raise your daily energy levels.

**2. Lower Blood Pressure**

Many green leafy vegetables like kale, spinach, and chard contain potassium. Potassium helps your kidneys filter sodium out of your body more efficiently, which can reduce your blood pressure.

**3. Lower Risk of Heart Disease**

Green leafy vegetables also contain vitamin K, which is believed to prevent calcium from building up in your arteries. This can lower your risk of arterial damage and help prevent many heart health complications in the future.

**4. Diabetes Control**

Vegetables are particularly high in fiber, which is needed for optimal digestion. They have a low glycemic index, so your blood sugar won’t rise quickly after a meal. The American Diabetes Association recommends at least 3 to 5 servings per day of non-starchy vegetables like broccoli, carrots, or cauliflower.

**5. Nutrition**

Vegetables are a rich source of folate, a B vitamin that helps your body make new red blood cells. Folate is especially important for children’s health and may also reduce the risk of cancer and depression.

The World Health Organization (WHO) recommends that adults eat at least 400g, or 5 portions, of fruit and vegetables (excluding potatoes, sweet potatoes and other starchy roots) per day to reduce the risk of disease. This number of fruits and vegetables also ensures adequate fiber intake and can also reduce total sugar intake. A national nutrition survey conducted by the Australian Government showed only 6.8% of Australians eat the recommended number of vegetables, whilst just over half (54%) met the recommendations for usual serves of fruit (Fruit and Vegetables, n.d.).

**Vegetable Prices**

Vegetables have the greatest supply and price volatility of any agricultural item. Vegetables are difficult to keep in a consistent supply and price because they are grown outside and their yields vary greatly depending on the weather. As a result, vegetables have a substantial economic impact. Despite the government’s best efforts, recurring weather shifts have generated instability in vegetable supply and price swings in recent years (Illankoon & Kumara, 2020).

Prices for vegetables fluctuate frequently, harming the financial interests of farmers, business owners, and consumers. Vegetable market prices are influenced by market supply and demand mechanisms, as well as a variety of internal and external influences. Weather fluctuations are one of these aspects, and they can have varied degrees of impact on the entire process of planting, harvesting, transporting, and selling vegetables. The supply chain for vegetables in China is more vulnerable to weather extremes due to the spatial concentration of vegetable production and the dispersed consumer base, resulting in an imbalance between supply and demand and abnormally large-scale fluctuations in vegetable prices that have an impact on people’s livelihoods (Yang et al., 2022).

Moreover, vegetable prices have an impact on farmers’ income, standard of living, and vegetable producing choices. This is due to the cyclical and seasonal swings that affect vegetable prices; the price trend will alter based on the season, demand, and other factors. Therefore, concerns about price stability have typically focused on government initiatives. Vegetable prices have varied sharply and frequently recently, having a variety of negative effects on supply-chain participants like farmers, transportation, wholesale, retail, and customers (Chen et al., 2018).

The quality of life of residents will be directly impacted by frequent increases in vegetable market prices, which will also bury hidden dangers for the fall in vegetable prices in the subsequent round of prices. In contrast, the continued decline in vegetable market prices results in prices that are too low and negatively impact the vital interests and production enthusiasm of vegetable farmers, leaving the supply of vegetables poorly protected and causing the subsequent round of skyrocketing prices. Regional vegetable price fluctuations have a significant negative impact on farmers and consumers, which has an impact on the macroeconomic stability and significantly lowers the degree of regional social welfare (IStudj on Vegetable Price Fluctuation and Its Impact in Guangxi - ProQuest, n.d.).

In the Philippines, sustainable soil nutrient-enhancing strategies involve the wise use and management of inorganic and organic nutrient sources in ecologically sound production systems (Janssen, 1993). The primary goal of integrated nutrient management (INM) is to combine old and new methods of nutrient management into ecologically sound and economically viable farming systems that utilize available organic and inorganic sources of nutrients in a judicious and efficient way. Integrated nutrient management optimizes all aspects of nutrient cycling. It attempts to achieve tight nutrient cycling with synchrony between nutrient demand by the crop and nutrient release in the soil, while minimizing losses through leaching, runoff, volatilization and immobilization (Jou and Hossner 1998). Providing higher economic returns per unit area and developing new export markets for high value crops in the Philippines has been identified as a priority by the Philippine Government and the Australian Centre for International Agricultural Research (ACIAR) as means of increasing economic growth and improving the standard of living of people living in rural areas. Regions VIII (Leyte), X (Northern Mindanao/Cagayan de Oro) and XI (Southern Mindanao/ Davao) have significant potential for expanding vegetable production. Moreover, they are seen as strategically important to the Australian Government, whereby efforts to improve the livelihoods of the populations in these areas could contribute to improving geo-political stability in the region (Tulin et al., 2023).

Pricing Theory and Law of Supply and Demand is used to interpret the gathered statistical data from previous studies in this systematic review of economics literature in the Philippines. It provides explanation on the responsiveness of seaweed farmers in the Philippines to the price changes in relation to the supply and demand that is changing from time to time. The strong demand of seaweed drives the market prices, and this may drive seaweed farmers to increase the seaweed production but when the price are low seaweed farmers have the tendency to leave their farms resulting in a low production of seaweed and in effect a decrease in seaweed supply. This problem may be enlightened using the law of supply and demand wherein when the demand is high — the price is high, and the supply is low. On the other hand, if the demand is low — the price is low, and the supply is high. As a result, the seaweed farmers are not responsive to price changes in the short-run and in the long-run (Guerrero, 2022).

**Forecasting Vegetable Prices**

Accurate agricultural price forecasting is critical to achieving sustainable and healthy agricultural development, and it is a hot research area in the agricultural industry. It explores traditional forecasting procedures, intelligent forecasting methods, and combination model forecasting methods, as well as the obstacles found in the current research landscape of agricultural commodity price prediction (Choong et al., 2021)

Farming is first and foremost defined by family work, which is limited by the availability of land, water, and capital resources. Choosing which agricultural items to produce must be made by farmers, however there are oftentimes not enough possibilities to enable the best farming decisions. Farmers must choose which vegetables will bring in the highest prices at harvest. The aforementioned issue was fixed by estimating pricing based on weather conditions using machine learning technologies. The hopeful results of the Prediction model are what have made Machine Learning so well-liked. This paper explores the use of multiple regression models to forecast vegetable prices, and its applicability has also been taken into account. In order to plan for their forthcoming crop and prevent hyperinflation, farmers benefit from being able to estimate vegetable prices (Kakulapati & Shaik, 2022).

Forecasting vegetable prices is critical in the agricultural industry for making sound judgments. This forecasting task is quite difficult. Because neural networks are self-adaptive, have a high learning capacity, and are adaptable, they are used to tackle a wide range of difficult tasks. This model forecasts the price of vegetables for the next day based on the previous price of time series data. This research compares and contrasts three machine learning algorithms: radial basis function, back propagation neural network, and genetically based neural network. The models are assessed, and the accuracy percentages show that the genetically based neural network outperforms propagation neural networks and radial basis functions in terms of vegetable price prediction accuracy (Article Detail - International Journal of Advanced Research, 2020).

In the Philippines, box-Jenkins technique and the Autoregressive Moving Average (ARMA) model were used to forecast onion production. Using historical data from the Philippine Statistics Authority, the ARMA (4,2) model was applied to construct an optimal forecasting solution. The model passed diagnostic tests with a mean absolute percentage error (MAPE) of 10.406%. The predicted yields for each quarter were highlighted, as were projections for onion output in 2023 and 2024. An examination of historical data indicated that weather trends, consumer demand, agricultural techniques, and imports and exports all contribute to periodic variations in onion supply. The study’s findings highlight the significance of employing exact forecasting models when deciding how to distribute resources, establish pricing, and place products on the market.

Crop climate calendars enhance traditional crop calendars by providing phenological states, cultivation techniques, and weather and climatic requirements that all crops must meet during a cropping season, as well as planting and harvest timings. The case for capturing this data in Benguet is compelling: the mountainous province benefits from the growth of high-value crops like as carrots, cabbage, and potatoes despite weather phenomena such as hail, frost, and various microclimates. The researchers conducted focus groups with municipal agriculturalists and farmer leaders in Atok, Benguet, to better understand their experiences and build their crop climate calendar. The calendars developed during this experiment could provide as a solid foundation for investigating the area’s climate-sensitive agricultural methods (Domingo, 2020).

# Chapter III

## RESEARCH METHODOLOGY

### Research Design

The research will be purely quantitative in nature. To be exact, the researchers will employ comparative time series forecasting (Hyndman & Athanasopoulos, 2021). In this case, the researchers will estimate Auto Regressive Integrated Moving Average (ARIMA) and Exponential Smoothing (ETS) models using the monthly vegetable prices from NVAT. The researchers will then compare the models and then choose the model with the higher accuracy and use it to forecast monthly vegetable prices.

### Locale of the Study

The study will be conducted in Bambang, Nueva Vizcaya, particularly at NVAT. The province of Nueva Vizcaya is strategically situated at the center of northern luzon. Please see [Figure 1](#fig-nv) for the map of Nueva Vizcaya. The NVAT and its location in the province is shown in [Figure 2](#fig-nvat). In the Cagayan Valley region, the province of Nueva Vizcaya is well-known for the NVAT. It contributes significantly to the local economy and acts as a vital hub for the trade of agricultural goods, specifically vegetables.

|  |
| --- |
| Figure 1: Map of Nueva Vizcaya (https://www.openstreetmap.org) |

|  |
| --- |
| Figure 2: Map of NVAT, Bambang, Nueva Vizcaya (https://www.google.com/maps) |

### Source and Subject of the Study

The study will focuses particularly on the available time series data of vegetable prices at NVAT.

### Data Gathering Procedure

The researchers will inquire at the NVAT office about its background and about its data policy. Moreover, Letters will be written to the NVSU President, the NVAT General Manager, and the Municipal Mayor of Bambang to ask permission to go to NVAT and gather data for the research (see Appendix A, B, and C respectively). The researchers will gather available time series data of vegetable prices. Throughout the data collection procedure, ethical considerations such as data confidentiality and approval from appropriate authorities will be strictly followed.

### Data Analysis and Treatment of Data

The process of forecasting monthly vegetable prices is illustrated in [Figure 3](#fig-workflow). The illustration is a modified version of the workflow introduced by Hyndman & Athanasopoulos (2021).

|  |
| --- |
| Figure 3: Data Treatment and Analysis Workflow |

The first step is to tidy or prepare the data in the correct format. This process involves loading the data into R, identifying missing values, and resolving missing values to come up with tidy time series of monthly vegetable prices. The researchers are going to use the tsibble and tidyverse packages for this task (Wang et al., 2020; Wickham et al., 2019).

The next step is to visualize the monthly time series of vegetable prices. Tidying and visualizing the data will allow the researchers to describe the data. The researchers will compute the maximum and minimum values and the time series of vegetable prices will be decomposed into their trend and seasonal components using the Seasonal and Trend decomposition using Loess (STL) method developed by Cleveland et al. (1990).

After this, the researchers will then fit compare the ARIMA and The next step fit the data to come up with ARIMA and

Data will be tidied into monthly time series of vegetable prices in R using the tidyverse package (Wickham et al., 2019). Tidied time series data will allow the data to be analyzed in R. The Researchers will split the data into training and testing data sets. The models will be specified using the the Using the fable package in R, ARIMA and ETC models will be fitted to the data. The ARIMA() and ETS() functions of the fable package have algorithms automatically choose the best ARIMA and ETS models

### Forecast Performance Evaluation

# References

Chen, J., Zhou, H., Hu, H., Song, Y., Gifu, D., Li, Y., & Huang, Y. (2018). Research on agricultural monitoring system based on convolutional neural network. *Future Generation Computer Systems*, *88*, 271–278. <https://doi.org/10.1016/j.future.2018.05.045>

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, *6*(1), 3–33. <http://bit.ly/stl1990>

*FAOSTAT*. (n.d.). Retrieved November 16, 2023, from <https://www.fao.org/faostat/en/?fbclid=IwAR2h4qJMCNFK7ryH25Q2IxT8e2SaZHwYrt73spPPQNcPg3FWx32VmqZ5iNA#data>

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts. [OTexts.com/fpp3](https://OTexts.com/fpp3)

Mchopa, A., Ruoja, C., & Huka, H. (2014). Price fluctuation of agricultural products and its impact on small scale farmers development: Case analysis from kilimanjaro tanzania. *European Journal of Business and Management*, *6*(36), 155–160. <https://catalog.ihsn.org/citations/91017?fbclid=IwAR06pCHPm9d6ko-uxXndkeBPGmpGgp7plCg0BQwlXNBHWh8OdHbY4B23soc>

Wang, E., Cook, D., & Hyndman, R. J. (2020). A new tidy data structure to support exploration and modeling of temporal data. *Journal of Computational and Graphical Statistics*, *29*(3), 466–478. <https://doi.org/10.1080/10618600.2019.1695624>

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. <https://doi.org/10.21105/joss.01686>