



Unveiling livestock trade trends: A beginner's guide to generative AI-powered visualization

Yoshiyasu Takefuji

Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku, Tokyo 135-8181, Japan

ARTICLE INFO

Keywords:

Generative AI
Livestock research
Data visualization
USDA datasets, non-programmers

ABSTRACT

This tutorial, rooted in the context of livestock research, is designed to assist novice or non-programmers in visualizing trends in livestock exports between the US and Japan using Python and generative AI systems such as Microsoft's Copilot and Google's Gemini. The analysis of these trends plays a pivotal role in optimizing livestock production. The tutorial offers a thorough guide on preparing data using reliable federal datasets, generating Python code, and tackling potential issues such as overlapping data points. It effectively simplifies complex tasks into manageable steps and includes Python code in the appendices for easy reference. By enabling researchers to extract insights and make predictions from livestock data, this tutorial addresses a significant void in the existing literature. This innovative approach has the potential to transform the way researchers engage with and interpret livestock data, thereby making a substantial contribution to the field.

1. Introduction

This tutorial is designed to introduce the application of generative AI in livestock research. It provides a practical guide on generating Python code to visualize trends in livestock import and export, using the most comprehensive and reliable datasets available. As an illustrative example, we will explore the export trends from the US to Japan. The data for this exercise is sourced from the official dataset provided by the Economic Research Service (ERS) at the U.S. Department of Agriculture (USDA). This tutorial is specifically tailored for livestock scientists, researchers, and practitioners who may not have prior programming experience. By making this guide accessible to a wider audience, we aim to bridge the gap between technology and livestock research.

This tutorial stands out as the first of its kind to offer a concise and comprehensive introduction to the use of generative AI in livestock science. It is our hope that this resource will empower livestock scientists to leverage the power of AI in their research.

While the key phrase "generative AI" was mentioned in 14 articles in this journal, there are currently no tutorial articles available. This paper is the first to address the application of generative AI in Veterinary Science research, aiming to revitalize and advance future research efforts using this innovative technology.

The paper emphasizes the importance of well-formed queries for generative AI systems. It highlights two free generative AI systems available to the public: Microsoft's Copilot with ChatGPT-4.0 (Morreel

et al., 2024; Takefuji, 2024) and Google's Gemini (Masalkhi et al., 2024; Mihalache et al., 2024). These systems play a crucial role in generating correct Python code for visualizing time-series trends.

Nazir et al. reviewed that Large Language Models (LLMs) combined with Generative Pre-trained Transformers (GPT) revolutionize natural language processing (Nazir and Wang, 2023). ChatGPT, a leading conversational model, excels in generating human-like responses. Their survey explored ChatGPT's inception, evolution, key technology, applications, and challenges, highlighting its significant impact on research and industry, while addressing ethical concerns and future research directions (Nazir and Wang, 2023).

Unlike the earlier paradigm of discriminative artificial intelligence, which mainly focuses on tasks like classifying or differentiating between classes in a dataset, generative AI models stand out by their ability to both discriminate and generate new information based on the input data (Tortora, 2024).

Fui-Hoon et al. explored generative AI's applications and challenges in healthcare, focusing on privacy issues, AI-human collaboration, and regulatory concerns (Fui-Hoon Nah et al., 2023). They highlighted risks like sensitive data disclosure, digital divide, data collection, copyright issues, lack of human control, data fragmentation, and information asymmetries between tech giants and regulators (Fui-Hoon Nah et al., 2023).

Gallifant et al. reviewed OpenAI's GPT-4 technical report, focusing on high-risk applications like healthcare (Gallifant et al., 2024). A

E-mail address: takefuji@keio.jp.

<https://doi.org/10.1016/j.rvsc.2024.105435>

Received 12 August 2024; Received in revised form 29 September 2024; Accepted 8 October 2024

Available online 10 October 2024

0034-5288/© 2024 Elsevier Ltd. All rights reserved, including those for text and data mining, AI training, and similar technologies.

diverse team analyzed the report, highlighting strengths such as transparent AI research and a comprehensive systems card. However, limitations include restricted training data access, inadequate confidence estimations, and privacy concerns. Their study recommended improving data transparency, accountability frameworks, and confidence standards, emphasizing the need for diverse, global involvement in LLM development and evaluation (Gallifant et al., 2024).

Yenduri et al. reported that the Generative Pre-trained Transformer (GPT) is a breakthrough in natural language processing, enabling machines to understand and communicate like humans (Yenduri et al., 2024). Based on transformer architecture, GPT excels in NLP tasks and has gained popularity among researchers and industries. Their review covered GPT's architecture, training, applications, challenges, and future directions, providing a comprehensive understanding of its impact and potential solutions (Yenduri et al., 2024).

OpenAI's release of ChatGPT has garnered significant attention from industry and academia due to its impressive capabilities (Zhou et al., 2024). Their large language model excelled in solving a variety of open tasks. They introduced its history, discussed its pros and cons, explored potential applications, and analyzed its impact on trustworthy AI, conversational search engines, and artificial general intelligence (Zhou et al., 2024).

Wach et al. identified challenges and opportunities of generative AI (GAI) in business, focusing on ChatGPT (Wach et al., 2023). They developed a conceptual framework through a narrative and critical literature review. Key findings included seven main threats: lack of regulation, poor quality control, job losses, data privacy issues, social manipulation, socio-economic inequalities, and AI technostress. Recommendations emphasized regulation, ethical considerations, and continuous skill development to mitigate these risks (Wach et al., 2023).

Kashefi et al. explored ChatGPT's capability for programming numerical algorithms, debugging, code completion, rewriting in different languages, and parallelizing serial codes (Kashefi and Mukerji, 2023). They assessed ChatGPT's performance on various mathematical problems and scientific machine learning applications. While ChatGPT showed success in programming numerical algorithms, it faces challenges like producing singular matrices and handling long codes, indicating areas for improvement (Kashefi and Mukerji, 2023).

By leveraging datasets, researchers can visualize trends and potentially predict future patterns using linear regressions. The paper encourages livestock scientists to explore time-series diverse trends and consider their implications for near-future predictions.

The visualization of livestock trends offers a wealth of opportunities for gaining insights, making predictions, and uncovering new findings. A thorough literature review has been conducted to examine existing research on trend visualization using livestock datasets. This review draws from peer-reviewed publications housed in the National Library of Medicine, providing a robust investigation into the intersection of data science and livestock studies. This approach ensures a comprehensive understanding of the current state of the field and the potential for future advancements.

Gilbert et al. introduced the Gridded Livestock of the World (GLW 3) database for various fields, including agricultural socio-economics, food security, environmental impact assessment, and epidemiology (Gilbert et al., 2018). It offered the most recent and harmonized subnational livestock distribution data for 2010, detailing global population densities of various livestock species. The data, presented in two versions, are disaggregated within census polygons either by statistical models or distributed homogeneously, providing spatial data layers free from assumptions linked to other spatial variables (Gilbert et al., 2018). However, they did not show programming and dataset future analysis.

Herrero et al. presented a unique, high-resolution global livestock dataset (Herrero et al., 2013). They provided comprehensive information on biomass use, production, feed efficiency, and greenhouse gas emissions across 28 regions and 8 livestock production systems. The dataset, which includes data for cattle, small ruminants, pigs, and

poultry, offers over 50 new global maps for understanding the multifaceted roles of livestock. It underscores the importance of feed efficiency, grasslands, and mixed crop-livestock systems in sustainable livestock production (Herrero et al., 2013). They did not discuss programming and future analysis.

Li et al. systematically reviewed the application of Convolutional Neural Network (CNN)-based computer vision systems in animal farming (Li et al., 2021). It explored five deep learning tasks: image classification, object detection, segmentation, pose estimation, and tracking, focusing on cattle, sheep/goats, pigs, and poultry. Their research summarized system development preparations, reviews CNN architectures, discussed algorithm development strategies, and evaluates model performance (Li et al., 2021). Their presentation was not for novice and non-programmers in livestock scientists.

Tuliozi et al. explored the genetic correlations of social dominance with various morphology and fitness traits in cattle (Tuliozi et al., 2023). Using a dataset of agonistic contests, it reveals that winning interactions correlate positively with developed frontal muscle mass, but negatively with fertility and udder health. Their study provided evidence of potential evolutionary trade-offs between these traits, suggesting that selection for social dominance may influence other aspects of livestock health and fitness (Tuliozi et al., 2023). However, they did not use generative AI to produce their results.

Moore studied that Japanese encephalitis virus (JEV) is a significant health concern in Asia, causing numerous cases of severe encephalitis and deaths annually (Moore, 2021). This study conducted a spatial analysis of risk factors associated with JEV, estimated the force of infection, and assessed the impact of vaccination from 2010 to 2019. It found that vaccination has significantly reduced the incidence of JEV, particularly in China, Taiwan, and Malaysia. The study underscored the importance of increasing vaccination coverage in at-risk populations to further reduce the burden of JEV (Moore, 2021).

While existing studies have not utilized generative AI in their visualization and analyses, there is a noticeable absence of tutorials on the application of generative AI in livestock research using datasets for novice and non-programmers. This paper presents an opportunity for the development of educational resources in this area to bridge the gap and enhance livestock research methodologies.

2. Materials and methods

While Generative AI may not yet be perfect at generating flawless Python code, it can be a valuable tool for beginners and non-programmers to visualize and analyze datasets. It's important for users to be familiar with CSV files as they often serve as the data source for visualizations involving multiple variables.

Before diving into Python code generation, it's crucial to understand the concept of task decomposition, which involves breaking down larger tasks into smaller, more manageable ones. This paper provides an example of this process by demonstrating how to visualize time-series export trends from the US to Japan.

The data for this visualization comes from a publicly available USDA file (USDA, 2024), which can be found online. The first step in the process is to download and unzip this file. This can be done programmatically, but it's recommended that readers manually download the `LivestockMeatTrade.zip` file and unzip it first. This allows them to see what CSV files are available after unzipping the file, providing a better understanding of the dataset they will be working with before they start making queries to the Generative AI.

Remember, understanding your data is a key step in any data analysis or visualization task. Once you have a good grasp of the data contained in these CSV files, you can then proceed to use Generative AI to assist with your Python code generation for data visualization. There are two csv files: `LivestockMeat_Exports.csv` and `LivestockMeat_Imports.csv`. Therefore, the first query to generative AI should be as follows.

Query1: Show Python code to download zip file from the following

and unzip it.

<https://www.ers.usda.gov/webdocs/DataFiles/81475/LivestockMeatTrade.zip>

Generative AI generates Python code and save the file named q1.py on the terminal which is shown in Appendix1. Run this code on the terminal and manually open LivestockMeat_Exports.csv file and make sure all ten variables in the first row such as SOURCE_ID, HS_CODE, COMMODITY_DESC, GEOGRAPHY_COD, GEOGRAPHY_DESC, ATTRIBUTE_DESC, UNIT_DESC, YEAR_I, TIMEPERIOD_ID, and AMOUNT.

We aim to visualize the total export amount and the total number of export instances per year from the US to Japan, using dataset from the `LivestockMeat_Exports.csv` file, which encompasses the time period from 2001 to 2024. Our program will provide an option to select a specific commodity, and subsequently generate a visualization that represents both the total export amount and the total number of exports for the selected commodity. The graph will be plotted with black dots representing data points, and the corresponding export counts will be printed in black text above each dot for clarity.

Query2: Show Python code to plot dots of export total amount with AMOUNT from the US to Japan with GEOGRAPHY_DESC from 2001 to 2024 with YEAR_ID using LivestockMeat_Exports.csv. Show possible candidates from COMMODITY_DESC determinant cell contents by number. User will enter number to select the single candidate from possible cell contents in COMMODITY_DESC. Visualize a graph of export total amount when two conditions are ANDed: GEOGRAPHY_DESC = "Japan" and COMMODITY_DESC = the selected commodity against the year (YEAR_ID). Print the total number of exported times per year on the top of plotted dots. DO NOT overlap printed black integer numbers and plotted black dots.

The generated Python code is attached in Appendix2. This code first loads the data from the CSV file into a pandas DataFrame. It then filters the data to include only rows where GEOGRAPHY_DESC is "Japan" and YEAR_ID is between 2001 and 2024. It then prints out the unique commodities and asks the user to select one. After the user has made a selection, it filters the data again to include only rows where COMMODITY_DESC is the selected commodity. It then groups the data by year and calculates the total export amount and the number of exports for each year. Finally, it plots the total export amount and prints the number of exports on top of the plotted dots. To avoid overlapping of the printed numbers and the plotted dots, a small random number is added to the y-coordinate of the text.

3. Results

Copilot with ChatGPT-4.0 was used as generative AI. To generate the results, run the following commands. You must install Python on your system before running the programs. (\$) sign indicates the prompt from the terminal.

3.1. \$ python q1.py

This generates two csv files such as LivestockMeat_Exports.csv and LivestockMeat_Imports.csv from zipped file, LivestockMeatTrade.zip.

3.2. \$ python q2.py

Figure 1 shows the outcome of q2.py when beef and veal was selected as commodity (number: 32) where Y-axis indicates the export total amount from the US to Japan. Note that the numerical value adjacent to each plotted dot signifies the total instances of exports for that year.

4. Discussion

If the resulting graph does not meet your expectations, feel free to provide feedback to the AI system about the current outcome and specify

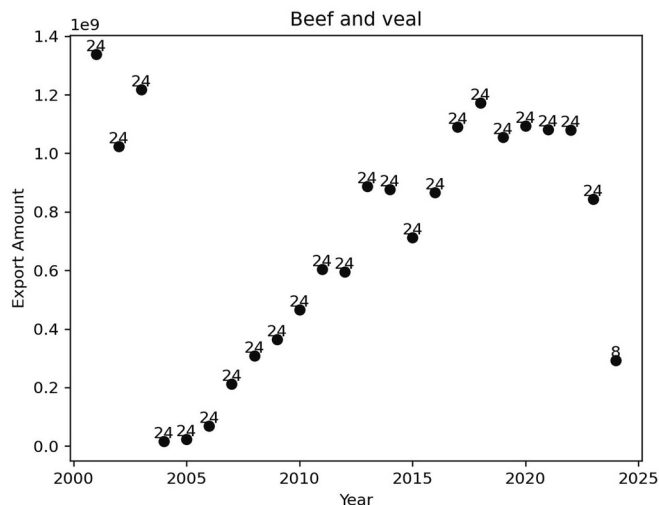


Fig. 1. Export of beef and veal from the US to Japan from 2001 to 2024.

the changes you'd like to see. A simpler query often yields better results with fewer requirements. Don't hesitate to add more requests or reiterate your requirements in detail for clarity. Generative AI models can't directly access or evaluate the outcome of their creations because they lack the ability to execute code or interact with the real world. This means you'll need to provide additional context or feedback about the results so the model can understand its effectiveness. The q2.py script has been enhanced to prevent overlapping integer values with plotted data points. Generative AI successfully implemented a new function to address this issue.

Generative AI offers customization for plot elements. Colored dots and numbers are the default choice when black text or black dots aren't specified. If you encounter errors in your generated Python code, try copying and pasting it into a query for generative AI. It may be able to suggest corrections or alternative approaches.

For reference, generative AI models typically achieve accuracy rates below 90 % when communicating in English. In other words, you may need to repeat or rephrase your requests or queries.

This tutorial has demonstrated the power of generative AI in making livestock research more accessible, particularly for those without prior programming experience. By leveraging generative AI tools like Copilot and Gemini, researchers can bridge the gap between complex data analysis and creating insightful visualizations. We have explored the process of generating Python code to visualize export trends from the US to Japan using a real-world USDA dataset. This exercise showcased how to break down complex tasks into smaller steps and utilize generative AI to automate code generation. The provided code offers functionalities such as user selection of specific commodities and addresses potential issues like overlapping data points. Remember, clear communication with generative AI systems is key to achieving desired results. Don't hesitate to refine your queries and provide feedback to the AI for optimal outcomes. This tutorial serves as a stepping stone for researchers to explore the vast potential of generative AI in livestock research. By embracing these new tools, researchers can unlock deeper insights from data, leading to more informed decision-making and advancements in the field.

Remember that generative AI is still imperfect, so multiple conversations may be needed to obtain the correct Python code for achieving the desired outcome. A step-by-step incremental approach is recommended for navigating generative AI.

Ethics in publishing

Not applicable.

Studies in humans and animals

Not applicable.

Funding

This research has no fund.

CRediT authorship contribution statement

Yoshiyasu Takefuji: Writing – review & editing, Writing – original

draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. q1.py—source code generated by Copilot with ChatGPT-4.0

```

Appendix 1: q1.py---source code generated by Copilot with ChatGPT-4.0
import requests
import zipfile
import os

# URL of the file to be downloaded
url = "https://www.ers.usda.gov/webdocs/DataFiles/81475/LivestockMeatTrade.zip"

# Send a HTTP request to the URL of the file
response = requests.get(url)

# Open the zip file in write binary mode
with open("LivestockMeatTrade.zip", "wb") as file:
    # Write the contents of the response to the file
    file.write(response.content)

# Path to the zip file
zip_path = "LivestockMeatTrade.zip"

# Directory to extract the zip file
extract_dir = "."

# Create a ZipFile object
with zipfile.ZipFile(zip_path, 'r') as zip_ref:
    # Extract all the contents of the zip file to the directory
    zip_ref.extractall(extract_dir)

print("The file has been downloaded and unzipped successfully!")

```

Appendix B. q2.py—Python code generated by generative AI

```

import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

# Load the data
df = pd.read_csv('LivestockMeat_Exports.csv')

# Filter the data
df = df[(df['GEOGRAPHY_DESC'] == 'Japan') & (df['YEAR_ID'].between(2001,
2024))]

# Show possible commodities
commodities = df['COMMODITY_DESC'].unique()
for i, commodity in enumerate(commodities):
    print(f'{i}: {commodity}')

# User selects the commodity
selected_index = int(input("Enter the number of the commodity: "))
selected_commodity = commodities[selected_index]

# Filter the data again
df = df[df['COMMODITY_DESC'] == selected_commodity]

# Group by year and calculate the total amount and count
grouped = df.groupby('YEAR_ID').agg({'AMOUNT': 'sum', 'COMMODITY_DESC':
'count'})

# Plot the total export amount
plt.scatter(grouped.index, grouped['AMOUNT'], label='Total Export Amount',
color='black')

# Print the total number of exported times per year on the top of plotted dots
for i, row in grouped.iterrows():
    plt.text(i, row['AMOUNT'] + np.random.randint(10, 18),
str(int(row['COMMODITY_DESC'])), color='black', ha='center', va='bottom')
    plt.title(f'{selected_commodity}')
    plt.xlabel('Year')
    plt.ylabel('Export Amount')
    plt.show()

```

Data availability

The author has no permission to share datasets.

References

- Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., Chen, L., 2023. Generative AI and ChatGPT: applications, challenges, and AI-human collaboration. *J. Inf. Technol. Case Appl. Res.* 25 (3), 277–304. <https://doi.org/10.1080/15228053.2023.2233814>.
- Gallifant, J., Fiske, A., Levites Strelakova, Y.A., Osorio-Valencia, J.S., Parke, R., Mwavu, R., Martinez, N., Gichoya, J.W., Ghassemi, M., Demner-Fushman, D., McCoy, L.G., Celi, L.A., Pierce, R., 2024. Peer review of GPT-4 technical report and systems card. *PLOS Digital Health* 3 (1), e0000417. <https://doi.org/10.1371/journal.pdig.0000417>.
- Gilbert, M., Nicolas, G., Cinardi, G., Van Boeckel, T.P., Vanwambeke, S.O., Wint, G.R.W., Robinson, T.P., 2018. Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Scientific data* 5, 180227. <https://doi.org/10.1038/sdata.2018.227>.
- Herrero, M., Havlik, P., Valin, H., Notenbaert, A., Rufino, M.C., Thornton, P.K., Blümmel, M., Weiss, F., Grace, D., Obersteiner, M., 2013. Biomass use, production, feed efficiencies, and greenhouse gas emissions from global livestock systems. *Proc. Natl. Acad. Sci. USA* 110 (52), 20888–20893. <https://doi.org/10.1073/pnas.1308149110>.
- Kashefi, A., Mukerji, T., 2023. ChatGPT for programming numerical methods. *J. Mach. Learn. Model. Comput.* 4 (2), 1–74. <https://doi.org/10.1615/JMachLearnModelComput.2023048492>.
- Li, G., Huang, Y., Chen, Z., Chesser Jr., G.D., Purswell, J.L., Linhoss, J., Zhao, Y., 2021. Practices and applications of convolutional neural network-based computer vision Systems in Animal Farming: a review. *Sensors (Basel, Switzerl.)* 21 (4), 1492. <https://doi.org/10.3390/s21041492>.
- Masalkhi, M., Ong, J., Waisberg, E., Lee, A.G., 2024. Google DeepMind's gemini AI versus ChatGPT: a comparative analysis in ophthalmology. *Eye (Lond.)* 38 (8), 1412–1417. <https://doi.org/10.1038/s41433-024-02958-w>.
- Mihalache, A., Grad, J., Patil, N.S., Huang, R.S., Popovic, M.M., Mallipatna, A., Kertes, P. J., Muni, R.H., 2024. Google Gemini and bard artificial intelligence chatbot performance in ophthalmology knowledge assessment. *Eye (Lond., Engl.)*. <https://doi.org/10.1038/s41433-024-03067-4>. Advance online publication.
- Moore, S.M., 2021. The current burden of Japanese encephalitis and the estimated impacts of vaccination: combining estimates of the spatial distribution and transmission intensity of a zoonotic pathogen. *PLoS Negl. Trop. Dis.* 15 (10), e0009385. <https://doi.org/10.1371/journal.pntd.0009385>.
- Morreel, S., Verhoeven, V., Mathysen, D., 2024. Microsoft Bing outperforms five other generative artificial intelligence chatbots in the Antwerp university multiple choice medical license exam. *PLOS Digital Health* 3 (2), e0000349. <https://doi.org/10.1371/journal.pdig.0000349>.
- Nazir, A., Wang, Z., 2023. A comprehensive survey of ChatGPT: advancements, applications, prospects, and challenges. *Meta-radiology* 1 (2), 100022. <https://doi.org/10.1016/j.metrad.2023.100022>.
- Takefuji, Y., 2024. Impact of COVID-19 on arthritis with generative AI. *Int. Immunopharmacol.* 133, 112032. <https://doi.org/10.1016/j.intimp.2024.112032>.
- Tortora, L., 2024. Beyond discrimination: generative AI applications and ethical challenges in forensic psychiatry. *Front. Psychol.* 15, 1346059. <https://doi.org/10.3389/fpsyg.2024.1346059>.
- Tuliozi, B., Mantovani, R., Schoepf, I., Tsuruta, S., Mancin, E., Sartori, C., 2023. Genetic correlations of direct and indirect genetic components of social dominance with

- fitness and morphology traits in cattle. *Genet. Select. Evol.*: GSE 55 (1), 84. <https://doi.org/10.1186/s12711-023-00845-8>.
- USDA, 2024. Accessed on June 18, 2024. <https://www.ers.usda.gov/webdocs/DataFiles/81475/LivestockMeatTrade.zip>.
- Wach, K., et al., 2023. The dark side of generative artificial intelligence: a critical analysis of controversies and risks of ChatGPT. *Entrepreneur. Bus. Econ. Rev.* 11 (2), 7–30. <https://doi.org/10.15678/EBER.2023.110201>.
- Yenduri, G., Ramalingam, M., Selvi, G.C., Supriya, Y., Srivastava, G., Maddikunta, P.K.R., Raj, G.D., Jhaveri, R.H., Prabadevi, B., Wang, W., Vasilakos, A.V., Gadekallu, T.R., 2024. GPT (generative pre-trained transformer)—a comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions. *IEEE Access* 12, 54608–54649. <https://doi.org/10.1109/ACCESS.2024.3389497>.
- Zhou, J., Ke, P., Qiu, X., et al., 2024. ChatGPT: potential, prospects, and limitations. *Front. Inf. Technol. Electron. Eng.* 25 (1), 6–11. <https://doi.org/10.1631/FITEE.2300089>.