Data-Driven Analysis of Employment Conditions and Food Security in the United States (2022)

Sidratun Nur Samia

Matriculation ID: 23020409

Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Abstract—This study explores the relationship between employment conditions and food security status in the United States using 2022 data. The datasets include employment details such as hours worked, gross pay, and job type, alongside household-level indicators of food insecurity. Using data pipelines, pre-processing, and analysis techniques, the data was integrated and analyzed to identify key insights. This paper aims to understand the socioeconomic factors contributing to food insecurity.

I. INTRODUCTION

Food security is a significant socioeconomic indicator that reflects household well-being and access to essential resources. Employment conditions, such as wages, working hours, and job stability, directly impact a household's ability to ensure food security. This study leverages publicly available datasets from the Medical Expenditure Panel Survey (MEPS) to examine how employment-related factors influence food security indicators.

II. DATA SOURCE

The analysis utilizes two primary datasets sourced from the **Medical Expenditure Panel Survey (MEPS)** for the year **2022**, provided by the **Agency for Healthcare Research and Quality (AHRQ)**. These datasets are essential for exploring the relationship between employment conditions and food security in the United States.

The **Job and Employment Conditions Dataset (HC-237)** [1] offers comprehensive information on employment characteristics, including factors such as employment type, hours worked per week, earnings, and the acceptance of employer-provided insurance benefits. Structured in a tabular format, each row corresponds to an individual respondent, while the columns represent key employment attributes. The dataset captures essential details such as unique household identifiers, job classification, weekly working hours, gross and daily wages, and whether insurance offered by employers was accepted. During preprocessing, the dataset underwent column standardization, renaming, and transformation to ensure numerical consistency across relevant fields. This ensured that any irregularities, such as missing or non-numeric values, were addressed efficiently.

The **Food Security Dataset (HC-240)** [2], on the other hand, focuses on household-level data concerning food security indicators. It provides insights into households' ability to afford balanced meals, the frequency of skipped meals due to financial constraints, and worries about running out of food. Like the employment dataset, the food security dataset is also presented in a tabular structure, with rows representing households and columns detailing food security metrics. The dataset includes information on whether households experienced anxiety about food supplies, if meals were skipped, and whether financial constraints prevented access to sufficient nutrition. Preprocessing ensured standardization of numeric fields and removal of invalid or missing entries to maintain data quality.

Both datasets underwent preprocessing steps, including data standardization, renaming of columns, and handling of missing or invalid values. The datasets were subsequently merged based on unique household identifiers to ensure alignment between employment characteristics and food security metrics.

A. Data Pipeline and Integration

The datasets were ingested, pre-processed, and merged to ensure consistency and reliability for analysis.

Figure 1 illustrates the **Data Pipeline Workflow**, including:

- **Data Ingestion:** Datasets were downloaded from publicly available sources.
- **Pre-processing:** Cleaning was performed to handle missing values and standardize data formats.
- **Merging Datasets:** Both datasets were merged using a common identifier (*dwelling_unit_id*).
- **Storage:** The final dataset was stored in an SQLite database and exported as a CSV file.
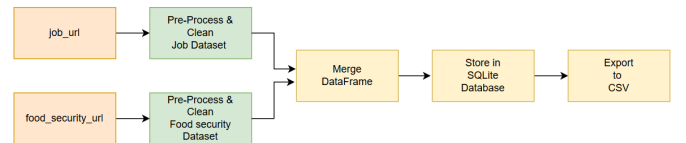

Fig. 1: Data Processing Pipeline

B. Merged Dataset Overview

The cleaned and pre-processed datasets were successfully merged, as shown in Figure 2. The merged dataset contains comprehensive information connecting employment characteristics and food security indicators.

Fig. 2: Merged Dataset Preview

C. Data Quality and License

Both datasets are sourced from the Medical Expenditure Panel Survey (MEPS) and are publicly available under an open-data license. They comply with transparency and accessibility standards and are suitable for academic analysis. Key considerations include:

- **Relevance:** The datasets provide detailed socioeconomic and employment insights.
- **Coverage Period:** The data spans the year 2022, ensuring temporal consistency.
- **Open Data License:** Both datasets are available for public use under AHRQ's open-access policy.

III. ANALYSIS

The analysis explores the relationship between employment conditions and household food security in the United States using the pre-processed and merged datasets. Various statistical and visualization techniques were employed to uncover patterns, trends, and correlations. The analysis focused on employment metrics such as gross pay, hours worked, and job type, as well as food security indicators such as worries about food running out, skipping meals, and meal affordability.

A. Data Overview

The merged dataset was loaded and subjected to an initial overview to understand its structure and quality. Key steps included generating summary statistics, checking data types, and identifying missing values. Summary statistics provided insights into the distribution of numeric fields, while missing value analysis ensured data completeness and consistency.

B. Distribution of Gross Pay

The distribution of gross pay was analyzed using a histogram (Figure 3). The visualization reveals that gross pay is highly skewed, with the majority of respondents earning relatively low salaries. Only a small percentage of individuals fall into higher salary brackets, indicating income disparity across the dataset.

C. Distribution of Hours Worked per Week

The analysis of hours worked per week (Figure 4) shows a strong concentration around the standard 40-hour work-week, with a few outliers working significantly more or fewer hours. This insight aligns with typical full-time employment patterns in the United States.

Fig. 3: Distribution of Gross Pay

Fig. 4: Distribution of Hours Worked Per Week

D. Food Security Indicators

Metrics related to food security, including worries about running out of food and skipping meals, were visualized to identify trends. Figure 5 shows that a significant portion of respondents expressed high levels of concern about food security, with a smaller proportion reporting meal skipping due to financial constraints.

Fig. 5: Worried About Food Running Out

In addition, skipped meals due to food insecurity were visualized (Figure 6), revealing a concerning number of households facing meal scarcity, further emphasizing the importance of addressing economic barriers to food access.

E. Gross Pay and Skipped Meals

The relationship between gross pay and meal skipping was analyzed using a boxplot (Figure 7). The results indi-
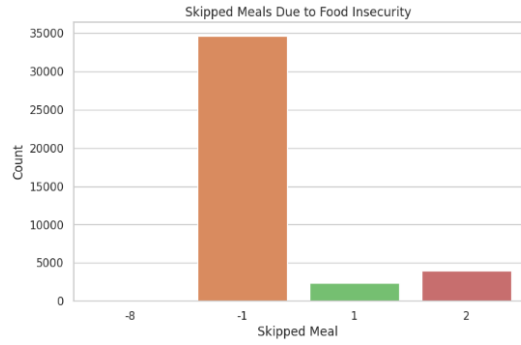
Fig. 6: Skipped Meals Due to Food Insecurity

cate that individuals with lower gross pay are more likely to skip meals due to financial constraints. Conversely, higher gross pay appears to reduce food insecurity.
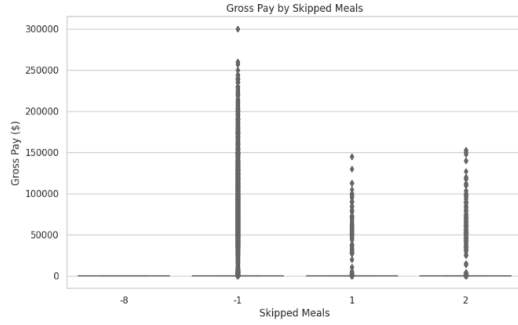

Fig. 7: Gross Pay by Skipped Meals

F. Hours Worked and Food Worry

A similar analysis was conducted to examine the relationship between hours worked per week and concerns about food running out (Figure 8). The results suggest that individuals working fewer hours per week are more likely to express concerns about food insecurity.
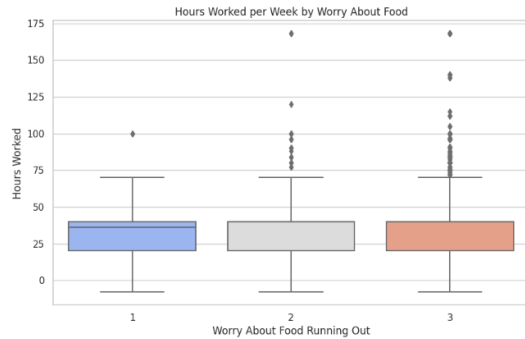
Fig. 8: Hours Worked per Week by Worry About Food

G. Correlation Analysis

A heatmap was generated to visualize the correlation between employment and food security variables (Figure 9).

Strong negative correlations were observed between gross pay and food insecurity indicators, while hours worked per week showed moderate associations with food-related concerns.
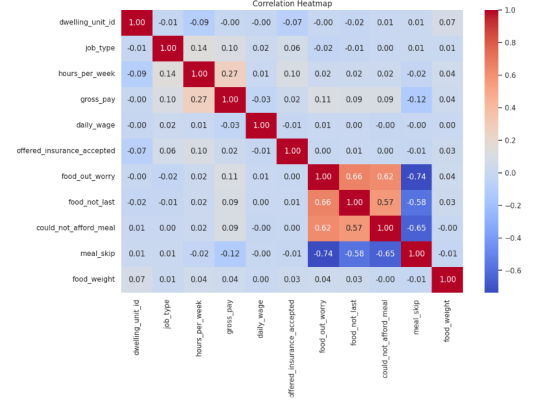
Fig. 9: Correlation Heatmap

H. Key Insights

The analysis yielded several important findings:

- Higher gross pay significantly reduces the likelihood of food insecurity, including concerns about running out of food and skipping meals.
- Individuals working fewer hours per week are more likely to experience food insecurity.
- Correlation analysis revealed strong associations between employment characteristics and household food security indicators.

These findings highlight the critical role of stable employment and adequate income in mitigating food insecurity and underscore the importance of targeted social policies to address these challenges.

A correlation heatmap highlighted significant relationships between employment and food security metrics, providing a visual representation of these dependencies.

IV. CONCLUSION

This study investigated the relationship between employment conditions and household food security in the United States using data from the Medical Expenditure Panel Survey (MEPS) for the year 2022. The analysis revealed significant relationships between employment metrics, such as hours worked per week, gross pay, and insurance acceptance, and key food security indicators, including meal skipping, concerns about food availability, and financial barriers to accessing balanced meals. Higher gross pay was found to correlate with lower levels of food insecurity, while fewer working hours were associated with increased concerns about running out of food. The insights derived from this analysis can serve as a foundation for policymakers to design targeted interventions aimed at reducing food insecurity and improving employment stability across vulnerable populations. Addressing

these socioeconomic factors can lead to more sustainable solutions for improving household welfare and reducing disparities in food access.

V. LIMITATIONS AND FUTURE WORK

Despite the robustness of the analysis, certain limitations must be acknowledged.

Firstly, **missing values** posed an initial challenge in the datasets. Some rows contained null or invalid entries, which were resolved by converting columns to numeric types and dropping rows with NaN values. Although this ensured cleaner data, it might have introduced biases by excluding certain records.

Secondly, **data merging issues** arose due to discrepancies in column naming conventions and data formats between the two datasets. This was addressed through column renaming and schema alignment, ensuring a seamless integration process. However, minor inconsistencies might still exist due to limitations in dataset granularity.

Additionally, the analysis is constrained by the static nature of the dataset, which represents a snapshot from 2022. The lack of real-time data limits the study's ability to capture dynamic trends or sudden economic shifts that might influence employment conditions and food security outcomes.

A. Data Quality Measures

To address these limitations, several quality control measures were implemented throughout the data pipeline:

- **Data Validation:** Columns were validated to ensure alignment with expected data types after preprocessing.
- **Error Handling:** Exception handling mechanisms were employed to manage issues such as failed data retrievals or unexpected data inconsistencies, ensuring the pipeline could log errors without crashing.
- **Adaptability:** The pipeline was designed to handle minor schema changes, such as additional or renamed columns, ensuring resilience against dataset updates.

B. Future Work

Future studies can address the identified limitations by:

- Incorporating **real-time datasets** to capture dynamic changes in employment and food security trends.
- Expanding the analysis to include **additional demographic variables** such as education level, geographic location, and household size to better understand context-specific influences.
- Utilizing **advanced machine learning models** to predict food insecurity outcomes based on employment metrics with higher accuracy.

In conclusion, while the current study provides valuable insights into the intersection of employment and food security, continued research is essential to refine these findings and address the evolving challenges faced by vulnerable populations.

REFERENCES

[1] *Medical Expenditure Panel Survey (MEPS) - HC-237 Job and Employment Conditions Dataset, 2022*. Available at: https://meps.ahrq.gov/mepsweb/data_files/pufs/h237/h237xlsx.zip

[2] *Medical Expenditure Panel Survey (MEPS) - HC-240 Food Security Dataset, 2022*. Available at: https://meps.ahrq.gov/mepsweb/data_files/pufs/h240/h240xlsx.zip