Charging Ahead: Socioeconomic Inequities in Electric Vehicle Charging Infrastructure Across the United States

Sagiv Antebi, Michael Fire
Department of Software and Information Systems Engineering Ben-Gurion University of the Negev Beer-Sheva, Israel
{sagivan, mickyfi}@post.bgu.ac.il

***Abstract*—The global shift towards sustainable transportation has fueled a rapid surge in electric vehicle (EV) adoption, creating an urgent need to understand how EV charging infrastructure is distributed across communities. This study investigates the distribution of EV charging stations across the United States and explores their relationship with key socioeconomic factors, including household income, population density, and land area. By merging datasets from the U.S. Department of Energy and Household Income Statistics, we conducted an exploratory data analysis (EDA) to investigate correlations between station count and income variables. We addressed data inconsistencies through rigorous cleaning and standardization processes, including city name normalization. Our findings reveal a positive correlation between the number of EV charging stations and mean household income, indicating that wealthier areas tend to have greater access to EV infrastructure. Temporal analysis shows an increasing trend in station openings, correlating with rising average incomes over the years. These insights emphasize the importance of addressing socioeconomic disparities to ensure equitable access to EV infrastructure, paving the way for more inclusive and sustainable transportation solutions.**

I. Introduction

The transportation sector is transforming with the increasing adoption of electric vehicles (EVs) as a sustainable alternative to traditional internal combustion engine vehicles. This shift is driven by the growing recognition of the environmental benefits of EVs, including the reduction of greenhouse gas emissions and the mitigation of air pollution. According to the International Energy Agency (IEA), global EV sales surpassed 10 million units in 2022, reflecting a robust compound annual growth rate that underscores the rapid penetration of EVs into mainstream markets [1].

In the United States, the adoption rate of EVs has been similarly promising, supported by federal incentives, advancements in battery technology, and expanding consumer awareness [2]. Central to the widespread adoption of EVs is the availability and accessibility of a comprehensive and reliable charging infrastructure. EV charging stations serve as critical enablers, alleviating range anxiety and the fear of depleting battery power without access to a charging point and ensuring that EV owners can conveniently recharge their vehicles during daily commutes, long-distance travel, and other routine activities [3].

However, the distribution of EV charging stations is not uniform across different geographic and socioeconomic landscapes. Socioeconomic factors, such as household income, population density, and land area, can significantly influence where charging stations are installed. Higher-income areas may exhibit greater demand for EVs and, consequently, more investment in charging infrastructure. Conversely, lower-income regions might face challenges in attracting such investments, potentially leading to disparities in access to EV charging services [4]. The disparities can worsen existing inequalities, limiting the benefits of EV adoption to more affluent communities and hindering equitable access to sustainable transportation.

Understanding the relationship between EV charging station locations and socioeconomic factors is essential for policymakers, urban planners, and stakeholders aiming to promote equitable and efficient deployment of EV infrastructure. By identifying patterns and correlations between charging station distribution and indicators of socioeconomic status, strategies can be developed to address gaps, optimize resource allocation, and ensure that the transition to electric mobility benefits diverse populations across the United States.

This study aims to analyze the spatial distribution of EV charging stations in the US and examine their interconnections with socioeconomic variables. Utilizing datasets from the U.S. Department of Energy and household income statistics, the research employs data cleaning, standardization, and exploratory data analysis techniques to prepare and merge the data effectively. Through correlation analysis and geospatial visualization, the study seeks to uncover insights into how socioeconomic factors influence the placement of EV charging infrastructure. The findings are intended to inform future efforts in creating a more equitable and accessible EV charging network, thereby supporting the broader goals of sustainable transportation and environmental stewardship.

II. Related Work

The deployment of electric vehicle (EV) charging infrastructure has been the focus of various studies aiming to understand the factors influencing its distribution and accessibility. Prior research has often concentrated on the technical and economic aspects of charging station implementation, such as optimal location planning based on traffic flow, grid capacity, and cost-effectiveness [5].

A significant contribution to the understanding of socioeconomic influences on EV infrastructure is the study titled *"Charging into Inequality: A National Study of Social, Economic, and Environmental Correlates of Electric Vehicle Charging Stations"* [6]. This national-level analysis revealed that EV charging stations are disproportionately situated in areas with higher income levels, greater racial homogeneity (predominantly white populations), and higher educational attainment. The authors argue that this unequal distribution may exacerbate existing social and economic disparities, limiting the benefits of EV adoption to more affluent communities.

Other studies have explored similar themes. For instance, Vergis and Chen [7] examined state-level factors affecting EV adoption in the United States, identifying household income and environmental policies as significant predictors. Similarly, Hardman et al. [8] analyzed consumer attitudes toward EVs, finding that financial incentives and charging infrastructure availability are critical for widespread adoption.

III. Data

A. Datasets Used

1) EV Charging Stations Data: The primary dataset for EV charging stations was obtained from the U.S. Department of Energy's Alternative Fuels Data Center, accessible at https://driveelectric.gov/stations. This dataset contains detailed information about EV charging stations across the US, including their geographic locations, operational status, and other relevant attributes. As of the time of data collection, the dataset comprised 73,725 entries with 75 attributes each, providing a comprehensive overview of the national charging infrastructure.

2) U.S. Household Income Data: Household income data was sourced from a publicly available dataset on Kaggle, titled *"US Household Income Statistics"* [9]. This dataset includes income statistics per city, such as mean, median, and standard deviation of household incomes, along with geographic identifiers, such as square area of land and water. The dataset covers 32,526 entries, representing various cities, towns, and villages across the United States.

3) Population Data: The population data for each state was obtained from Wikipedia, titled *"List of U.S. states and territories by population"* [10]. This dataset provides population statistics for all U.S. states and territories. The data was processed using GPT-4o to extract the relevant numbers, which were rounded to simplify the analysis.

B. Data Preparation and Cleaning

1) EV Charging Stations Data: Initial inspection of the EV charging stations dataset revealed several issues that required cleaning:

- **Removal of Empty Columns:** Columns with entirely missing values were identified and removed. Specifically, 33 columns containing no data were dropped to streamline the dataset.
- **Handling Missing and Duplicate Entries:** Rows with missing values in critical fields such as *City* and *State* were examined. Duplicate entries were also identified based on the combination of *State* and *City* fields.
- **Selection of Relevant Columns:** For the purposes of this study, only essential columns were retained: *ID*, *Station Name*, *Street Address*, *City*, *State*, *ZIP*, and *Open Date*.

After cleaning, the dataset was grouped by *State* and *City*, aggregating other fields such as *Station Name*, *Street Address*, and *Open Date* into lists. This resulted in a dataset with 8,772 unique city-state combinations. The number of stations per city was calculated, revealing a skewed distribution with a few cities having a large number of stations and many cities with only one station.

2) Household Income Data: The household income dataset required similar cleaning steps:

- **Renaming Columns:** The *State_ab* column was renamed to *State* for consistency with the EV charging stations dataset.
- **Handling Duplicates and Missing Values:** Duplicate entries were checked based on the *State* and *City* fields. Missing values were minimal and did not significantly impact the dataset.
- **Aggregation:** The dataset was grouped by *State* and *City*, aggregating income statistics (*Mean*, *Median*, *Stdev*) by calculating their means. Geographic area fields (*ALand* and *AWater*) were summed to represent the total land and water area for each city.

After grouping, the household income dataset consisted of 11,228 unique city-state combinations.

C. Standardizing City Names

A critical challenge encountered during the data merging process was the inconsistency of city names between the two datasets. Variations in city naming conventions, such as "Saint Louis" vs. "St. Louis" or abbreviations and alternate spellings, hindered the accurate merging of records.

To address this issue, a comprehensive city name standardization script was developed. The script employed the following strategies:

- **Static Replacements:** A dictionary of common city name variations was created using GPT-4o to map known aliases to a standardized form (e.g., "Saint Louis" mapped to "St. Louis").
- **Pattern-Based Substitutions:** Regular expressions were used to handle common prefixes and suffixes, such as replacing "Mount" with "Mt.", "Fort" with "Ft.", and directional words like "North" with "N".
- **Lowercasing and Cleaning:** All city names were converted to lowercase, and extraneous characters such as periods and extra spaces were removed to ensure uniformity.

The standardization process improved the overlap between the two datasets, increasing the number of matching city-state combinations from 5,475 to 5,939.

D. Data Merging

With standardized city names, the EV charging stations dataset and the household income dataset were merged on the *State* and *City_standardized* fields. The merged dataset included:

- **EV Charging Stations Information:** Total station count per city, station names, street addresses, and open dates.
- **Household Income Statistics:** Mean, median, and standard deviation of household incomes, land area (*ALand*), and water area (*AWater*).

As of 2018, there are 19,495 incorporated cities, towns, and villages in the United States, with 14,768 of these having populations below 5,000. The merging process resulted in 5,939 combined records of cities in the U.S., a significant portion considering the distribution of city populations. This highlights the completeness of the dataset and ensures its suitability for subsequent analysis. The data about the number of U.S. cities was referenced from [11].

E. Exploratory Data Analysis

Exploratory data analysis (EDA) was conducted to understand the distributions and relationships within the merged dataset.

1) Normalization and Visualization: Continuous variables such as *Mean*, *Median*, *Stdev*, *ALand*, and *AWater* were normalized using min-max scaling for consistent scaling in visualizations. Histograms and box plots were generated to examine the distributions, revealing the presence of outliers and skewness in certain variables.

2) Correlation Analysis: Pearson correlation coefficients were calculated between *Station Count* and other variables. Notably, a modest positive correlation was observed between *Station Count* and *Mean* household income ($r = 0.054$). This finding suggests that cities with higher average incomes do not necessarily have more charging stations. The apparent lack of a stronger correlation will be further explored, particularly considering the potential bias introduced by not normalizing using population size.

3) Outlier Removal: To improve the robustness of the analysis, outliers were identified and removed using the Interquartile Range (IQR) method.

4) Temporal Analysis: An analysis of the number of stations opened over time was performed using the *Open Date* field. The dates were cleaned and standardized, and the number of stations opened each year was calculated. A significant increase in station openings was observed in recent years, aligning with the growing adoption of EVs.

5) State-Level Distribution of EV Charging Stations: Figure 2 illustrates the absolute distribution of EV charging stations by state. California, Texas, and New York lead in total station count, reflecting their economic strength and progressive adoption of EV policies. Income levels, represented by the color gradient, align closely with the distribution of charging stations, emphasizing the socioeconomic disparities in infrastructure deployment.

6) Mean Household Income Per State: The mean household income across states is visualized in Figure 3. Wealthier states, such as California and Massachusetts, exhibit higher income levels, correlating with higher densities of EV charging stations. This reinforces the idea that infrastructure deployment is closely tied to economic resources.

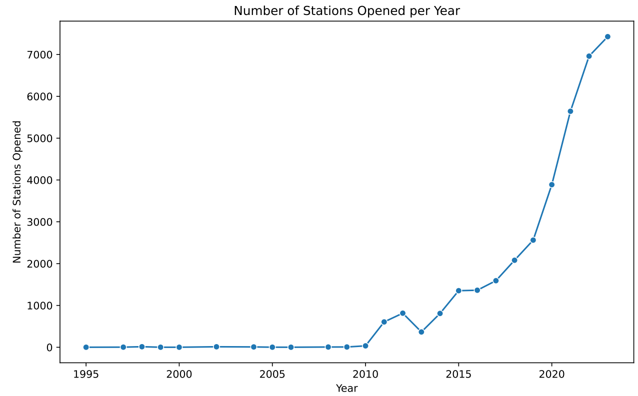

Fig. 1. Temporal trend in charging station openings over time, highlighting the evolution of infrastructure development.

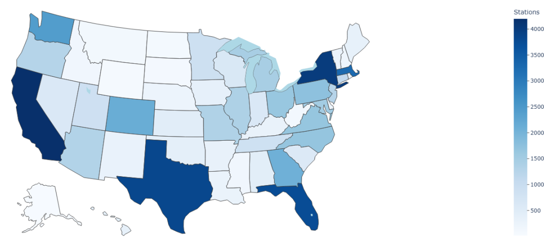

Fig. 2. Distribution of EV charging stations by state. The color gradient represents income levels, with darker shades indicating higher income.

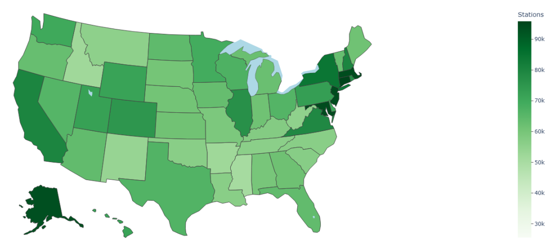

Fig. 3. Mean household income per state. The color gradient represents income levels, with darker shades indicating higher income.

IV. Results

A. City-Level Correlation Analysis

An initial analysis was conducted to examine the correlation between the number of EV charging stations and socioeconomic variables, such as mean and median income, land area, and water area, on a city-level basis. The single-row heatmap presented in Figure 4 reveals a weak relationship between the number of EV charging stations and the other features. The correlation coefficient between the number of stations and mean income is extremely low, indicating that income alone is not a strong predictor of EV infrastructure deployment when analyzing individual cities.

Given the lack of significant correlations at the city level, we decided to aggregate the data at the state level. By averaging income and summing land and water area for each state, we aimed to investigate whether correlations become more apparent when considering the larger scale of states instead of individual cities.

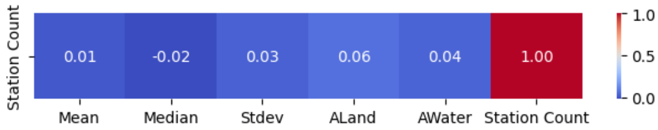

Fig. 4. City-level single-row heatmap between the number of EV charging stations, mean income, land area, water area, and other features.

B. State-Level Correlation Analysis

After aggregating city-level data into state-level data, the correlation between the features and the number of EV charging stations was re-evaluated. The single-row heatmap in Figure 5 demonstrates a strong correlation between population and the number of stations ($r = 0.88$). This finding indicates that population size is a key driver of EV infrastructure deployment at the state level.

Additionally, a moderate positive correlation ($r = 0.36$) is observed between mean income and the number of stations, suggesting that income plays a role in influencing infrastructure distribution but is less significant than population.

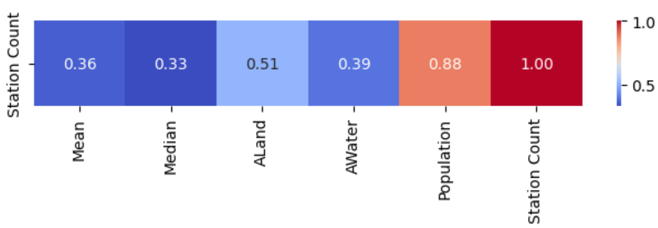

Fig. 5. State-level single-row heatmap between the number of EV charging stations, population, and mean income.

C. Geospatial Correlation Between EV Charging Stations and Household Income

The distribution of EV charging stations across the United States reveals a notable relationship with household income levels. As illustrated in Figure 6, wealthier regions tend to host more EV charging stations. The intensity of the color gradient represents mean household income, with brighter shades indicating higher income, while bubble sizes reflect the number of stations in each state. In Figure 6, the light green regions, representing the top 30% of income levels, clearly demonstrate the influence of income on the deployment of EV charging infrastructure. However, these results are incomplete as they do not account for the potential bias introduced by population size, which may disproportionately impact the observed distribution.

1) Normalized EV Charging Stations by Population: After normalizing the number of EV charging stations by population size, the distribution becomes more evident (Figure 7). The bright green and bright yellow regions, representing the top 30% and top 10% of income levels, respectively, have the largest circles. This indicates that high-income areas tend to have significantly better per capita access to EV charging infrastructure, even when accounting for population size.

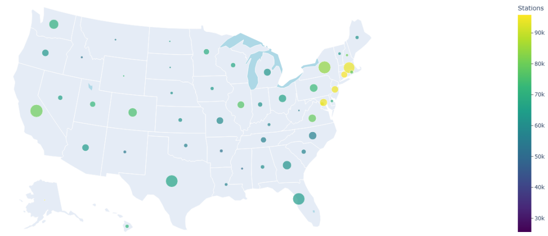

Fig. 6. Correlation between EV charging stations and household income across the United States. The color gradient represents income levels, and bubble size indicates the number of stations.

The population data for each state was sourced from Wikipedia [10] and processed using GPT-4o. Relevant population values were extracted, and numbers were rounded for simplicity.

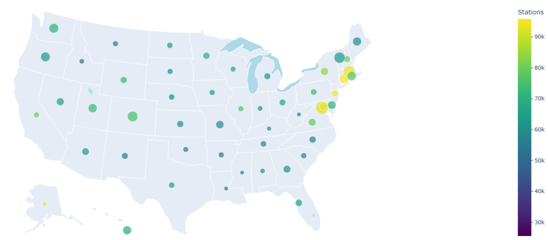

Fig. 7. Correlation between household income and EV charging stations normalized by population. The color gradient represents income levels, and bubble size indicates the normalized station count.

2) Normalized EV Charging Stations by Land Area: After normalizing by population, it is also important to normalize by land area to account for the geographic size of states, as larger states might require more infrastructure to ensure equitable access (Figure 8). The bright yellow regions, representing the top 10% of income levels, have the largest circles. This clearly indicates that wealthier states not only have higher income levels but also significantly better station density per land area, highlighting the strong influence of economic factors over geographic coverage needs.

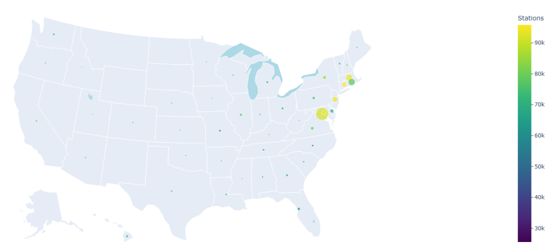

Fig. 8. Correlation between household income and EV charging stations normalized by land area. The color gradient represents income levels, and bubble size indicates the normalized station count.

D. Population vs. Number of EV Charging Stations with Correlation to Average Household Income per State

Figure 9 shows a clear linear trend, highlighting a strong correlation between population and the number of stations. Notably, states like *Massachusetts, New York, and Washington* stand out as outliers, with significantly more stations than expected, likely due to their large populations and advanced infrastructure efforts.

The color gradient, representing income levels, reveals that brighter circles (indicating higher-income states) tend to appear above the line. This suggests that wealthier states have more stations than expected for their population size, reinforcing the influence of income on EV infrastructure deployment.

Fig. 9. Population vs. number of EV charging stations across all states. Brighter circles represent higher-income states, while the line indicates the linear trend.

Focusing on states with populations under 5 million, Figure 10 provides a closer look at the relationship between population and station count. The same trend is visible here, with brighter circles (higher-income states) appearing above the line. This pattern suggests that even among smaller states, income plays a crucial role in determining the availability of EV infrastructure. States like *Connecticut and Maryland* outperform others with similar populations, further illustrating how income influences station deployment.

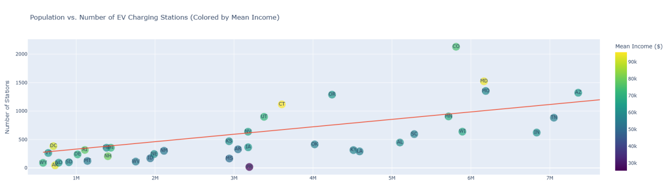

Fig. 10. Population vs. number of EV charging stations for states with populations under 5 million. Brighter circles represent higher-income states, while the line indicates the linear trend.

E. Temporal Correlation Between Income and Station Openings

An additional analysis was conducted to explore the temporal correlation between mean household income and the number of EV charging stations opened per state over the years. Figure 11 illustrates the evolution of this correlation from 2005 to 2024.

The results reveal three distinct phases: between 2005 and 2007, a strong positive correlation is observed, indicating that wealthier states disproportionately benefited from early EV infrastructure deployment. However, from 2008 to 2010, the correlation sharply declines and even turns negative, likely reflecting the impact of the global financial crisis, which may have disrupted infrastructure investment patterns and priorities across states.

From 2010 onward, the correlation stabilizes at approximately 0.25, indicating a modest but consistent relationship between mean income and station openings. This steady trend suggests that while wealthier states continue to have an advantage in station deployment, other factors, such as population size or state-level policies, may now play a more significant role in influencing EV infrastructure growth. These findings emphasize the evolving dynamics of socioeconomic factors in shaping the accessibility of EV infrastructure over time.

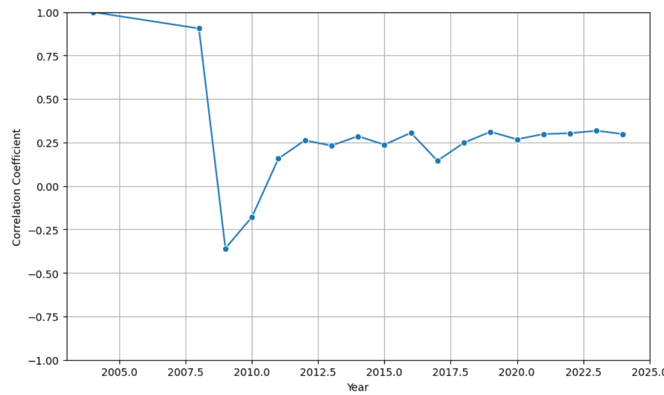

Fig. 11. Correlation between mean household income and EV charging stations opened per state over years. The strong correlation from 2005 to 2007, the dip from 2008 to 2010, and the stabilization around 0.25 from 2010 onward highlight the changing influence of income on infrastructure deployment.

V. Conclusion and Future Work

In this study, we investigated the socioeconomic inequities in the distribution of EV charging infrastructure across the United States. Our analysis revealed that higher-income states are significantly better equipped with EV charging stations, while economically disadvantaged areas face substantial barriers to access. The observed correlation between household income and charging station availability underscores the critical role of economic factors in shaping access to sustainable transportation solutions. These disparities highlight an urgent need to address inequities in the distribution of EV infrastructure to ensure a more inclusive transition to electric mobility.

Future work should focus on expanding the analysis to include additional factors such as racial demographics, urban versus rural divides, and the influence of local government policies on infrastructure deployment. Furthermore, developing predictive models to identify underserved regions could guide policymakers in prioritizing investments. Efforts should also be directed toward designing strategies that normalize infrastructure planning using population size and geographic area to better capture regional needs.

Addressing these inequities is essential for ensuring that the benefits of electric vehicle adoption are accessible to all communities, irrespective of socioeconomic status. Proactive and data-driven approaches will be instrumental in promoting an equitable and sustainable transportation future, bridging the gap between diverse communities, and fostering environmental and social justice.

REFERENCES

[1] International Energy Agency, "Global ev outlook 2023," https://www.iea.org/reports/global-ev-outlook-2023, 2023.

[2] U.S. Department of Energy, "Electric vehicle benefits and considerations," https://afdc.energy.gov/fuels/electricity-benefits, 2023.

[3] B. K. Sovacool and R. F. Hirsh, "Beyond batteries: An examination of the benefits and barriers to plug-in hybrid and battery electric vehicles," *Energy Policy*, vol. 148, p. 111912, 2020.

[4] S. Guo and E. Kontou, "Disparities and equity issues in electric vehicles rebate allocation," *Energy Policy*, vol. 154, p. 112291, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301421521001609

[5] S. He, Y. Yin, and J. Zhou, "Optimal deployment of public charging stations for plug-in hybrid electric vehicles," *Transportation Research Part B: Methodological*, vol. 47, pp. 87–101, 2013.

[6] B. K. Sovacool, J. Kester, L. Noel, and G. Z. de Rubens, "Charging into inequality: A national study of social, economic, and environmental correlates of electric vehicle charging stations," *Energy Research & Social Science*, vol. 44, pp. 206–218, 2018.

[7] S. Vergis and B. Chen, "Comparison of plug-in electric vehicle adoption in the united states: A state by state approach," *Research in Transportation Economics*, vol. 52, pp. 56–64, 2015.

[8] S. Hardman, A. Jenn, and G. Tal, "The effectiveness of financial purchase incentives for battery electric vehicles – a review of the evidence," *Renewable and Sustainable Energy Reviews*, vol. 80, pp. 1100–1111, 2017.

[9] Kaggle, "Us household income statistics with geographical information," https://www.kaggle.com/datasets/goldenoakresearch/us-household-income-stats-geo-locations, 2018.

[10] Wikipedia contributors, "List of U.S. states and territories by population," 2024, accessed: 2024-12-04. [Online]. Available: https://en.wikipedia.org/wiki/List_of_U.S._states_and_territories_by_population

[11] World Population Review, "How Many Cities are in the United States? 2024," 2024, accessed: 2024-12-05. [Online]. Available: https://worldpopulationreview.com/us-city-rankings/how-many-cities-are-in-the-us