

TEAM: RIO

PROJECT TITLE

KANSAS BROADBAND ACCESS AND USAGE

ANALYSIS REPORT

Developed by:

University Name: Wichita State University

In association with:

Yan, Fujian, Wichita State University

May 07, 2024

Contents

Summary	2
Objective	2
Introduction	3
Preparation	3
Analysis	4
Results	13
Recommendations	14
Conclusion	16

Summary

This report presents a comprehensive analysis of broadband access and usage in Kansas, comparing it against national averages and other states like California. Through a series of data visualizations, we have explored trends, disparities, and progress in broadband connectivity from 2019 to 2020, offering insights to inform the Kansas Office of Broadband Development's strategies for enhancing digital equity.

The Kansas Office of Broadband Development understands the essential role that digital equality plays in ensuring that all residents have equitable access to and opportunities in the digital age. The office has created a thorough Digital Equity Plan to address the growing demand for fair access to broadband services and digital resources. The objective of this study is twofold: firstly, to gain insights into the current state of broadband connectivity in different states, and secondly, to identify factors that influence broadband adoption and usage within these states and benchmarking them against Kansas. By understanding the factors driving regional differences in broadband usage, policymakers can devise targeted interventions to bridge the digital divide and promote digital inclusion.

Objective

The Kansas Office of Broadband Development is tasked with constructing a digital equity plan as part of a program to receive millions of federal dollars to increase broadband access in the state. They have asked us to provide data insights on broadband accessing the state of Kansas and benchmark it against other states.

Main Objectives of this data analysis project are:

- Understand internet data usage patterns across states by exploring the provided datasets.
- Identify trends and variations in internet speed and data usage over time.
- Develop dashboards to compare performance metrics, such as internet speed and data usage, among different States through visual analysis.

Introduction

The Kansas Office of Broadband Development aims to increase broadband access across the state, leveraging federal funds to ensure that residents and businesses benefit from high-speed internet connectivity. This report examines broadband usage and availability within Kansas, benchmarks it against other states, and provides insights into it, aiming to support the formulation of an effective digital equity plan.

In this exploratory data analysis (EDA) project, we focus on examining and comparing broadband usage across various states within the United States. By analyzing a comprehensive dataset sourced from Microsoft and the Federal Communications Commission (FCC), we aim to uncover regional variations in broadband adoption and usage patterns.

The objective of this study is twofold: firstly, to gain insights into the current state of broadband connectivity in different states, and secondly, to identify factors that influence broadband adoption and usage within these states. By understanding the factors driving regional differences in broadband usage, policymakers can devise targeted interventions to bridge the digital divide and promote digital inclusion.

Preparation

Data Overview

The analysis began with data examination and cleaning, focusing on datasets detailing broadband data for 2019, 2020 and by zip code. We filtered data to focus on Kansas, ensuring a clean, consistent dataset for in-depth analysis. Key steps included:

- Loading and examining datasets for understanding structure and content.
- Cleaning data to address any inconsistencies or missing values, specifically focusing on Kansas.

Datasets available: Currently we have two datasets

- FCC(Federal Communication commission): broadband_data_2019November, broadband_data_2020October & broadband_data_zipcode
- UK Data

Preprocessing

Data Collection

The analysis utilizes two primary datasets: `broadband_data_zipcode`, detailing broadband usage at the zip code level providing county-level broadband availability and usage statistics for 2019 and 2020.

Data Cleaning: No missing and duplicate values were found in FCC datasets, indicating good data completeness.

Data Integration: Three datasets from FCC were joined on `ST` and `COUNTY_ID` which are common data columns among them for further analysis.

Exploratory Data Analysis (EDA)

Distribution Analysis: The distribution of broadband usage rates 2020 appears to be somewhat right skewed, with a higher concentration of counties having lower usage rates. This suggests that while some areas have high broadband adoption, a significant number of counties experience lower usage levels.

The Right histogram shows a distribution of broadband usage rates that is right skewed towards lower rates, with a notable peak in the lower rate intervals. The height of the bars indicates that a significant portion of counties have broadband usage rates between approximately 10% and 30%.

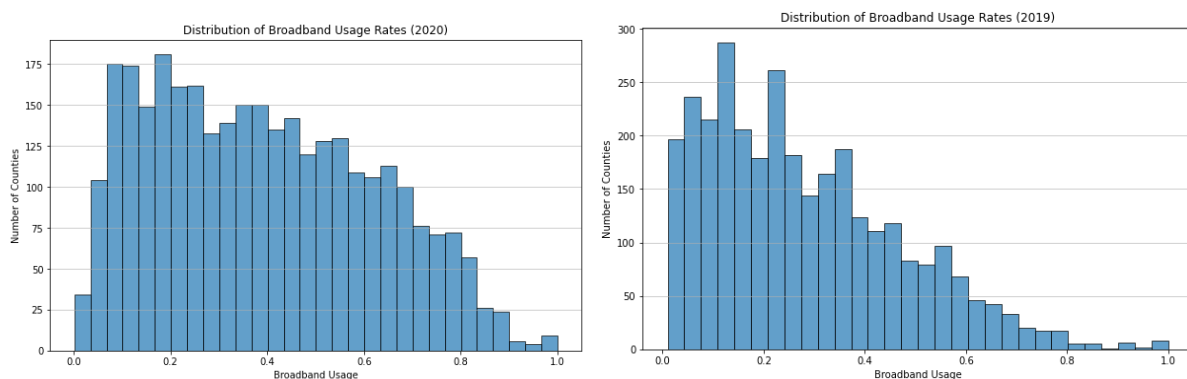


Figure 1: Distribution plots for Broadband Usage

Scatterplot Analysis:

- **Kansas's Positioning:** In 2019, Kansas's position, indicated by the red marker, shows its broadband availability and usage relative to other states. Like the 2020 data, Kansas appears to be within the middle range of states in terms of both metrics.
- **Nationwide Trends:** The plot suggests a positive correlation between broadband availability and usage across states, underscoring the importance of access as a driver of digital engagement.

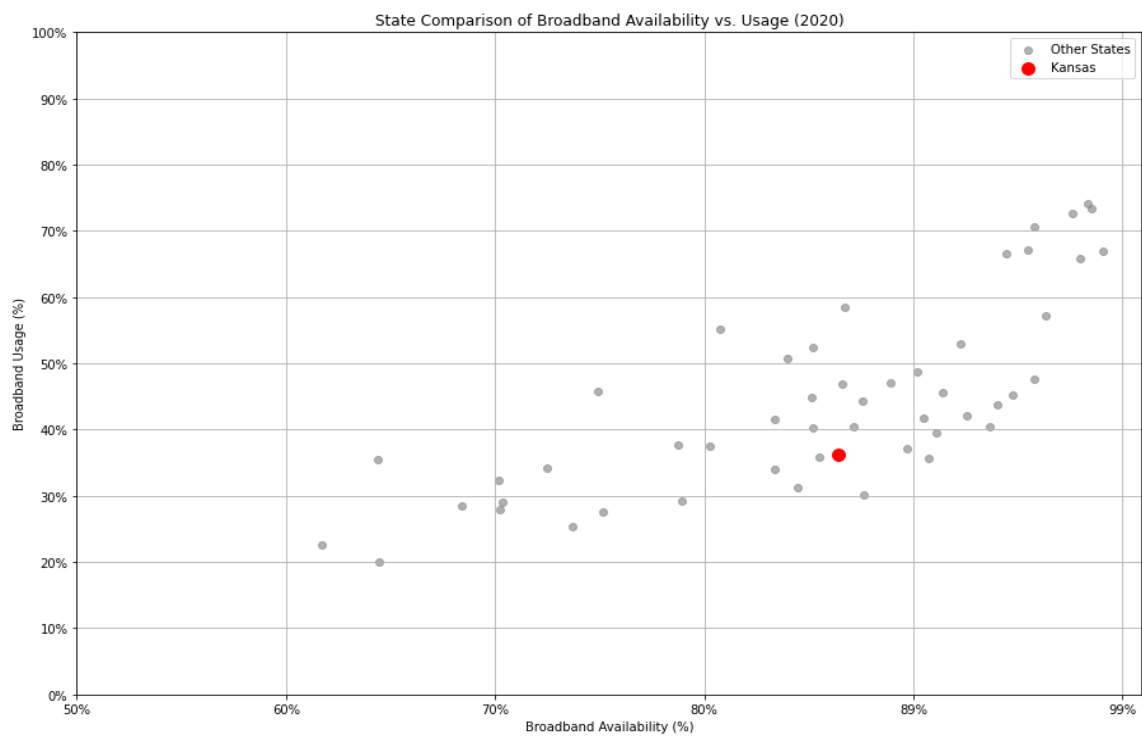
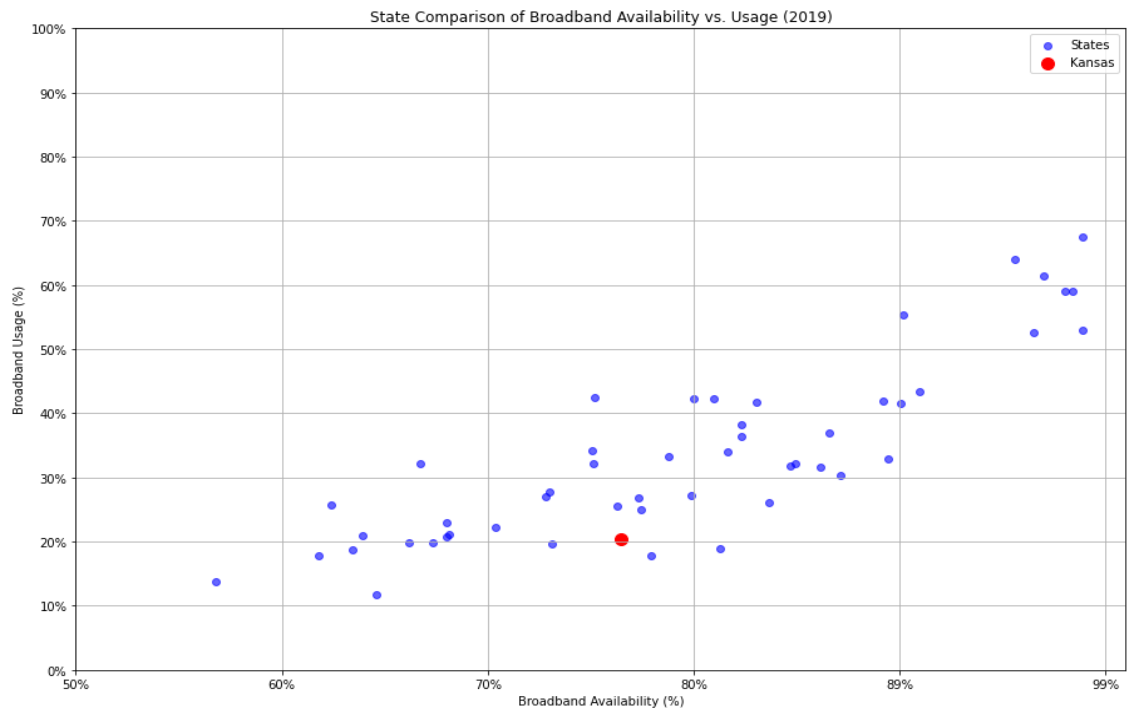


Figure 2: Scatter plots for Broadband Availability VS Usage

- **Kansas's Position:** Kansas is positioned well within the mix of states, showing a balanced improvement in both availability and usage. The red marker indicates Kansas's specific location based on the 2020 data.
- **General Trend:** There seems to be a positive correlation between broadband availability and usage among states, which is expected. States with higher availability typically have higher usage rates.

Broadband Availability Usage and Comparison:

- **Broadband Availability Comparison:** The first chart shows how Kansas compares with California (CA), New York (NY), and Texas (TX) in terms of broadband availability. This metric indicates the percentage of the population that has access to broadband services, reflecting the infrastructure in place.
- **Broadband Usage Comparison:** The second chart focuses on broadband usage, which reflects the percentage of the population actually utilizing broadband services. This metric is influenced by factors such as affordability, digital literacy, and the perceived value of internet services.

Insights:

- The charts can provide insights into how Kansas stands in terms of both broadband infrastructure (availability) and the effectiveness of its utilization (usage) compared to other states.
- If Kansas's bars are lower than those of other states, it might indicate areas for improvement in terms of infrastructure development and initiatives aimed at increasing broadband adoption.
- Conversely, if Kansas compares favourably, it may highlight successful policies or conditions that facilitate broadband access and adoption.

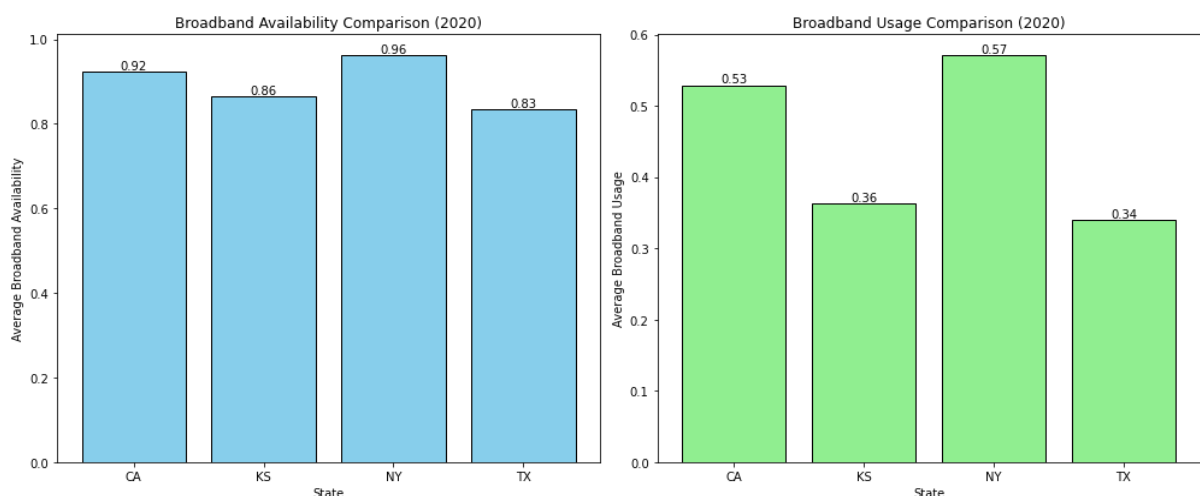


Figure 3: Broadband Usage and Availability Comparison

Average Broadband Usage by State (2020)

The map of the United States showing average broadband usage by state in the year 2020. The color gradient represents the percentage of broadband usage, with darker shades indicating higher usage and lighter shades representing lower usage.

Areas such as the Northeast and parts of the West Coast appear to have higher usage rates, whereas some central and southeastern states have lower usage. This could be indicative of various socio-economic factors such as urbanization, income levels, or state-level investment in broadband infrastructure.

Average Broadband Usage by State (2020)

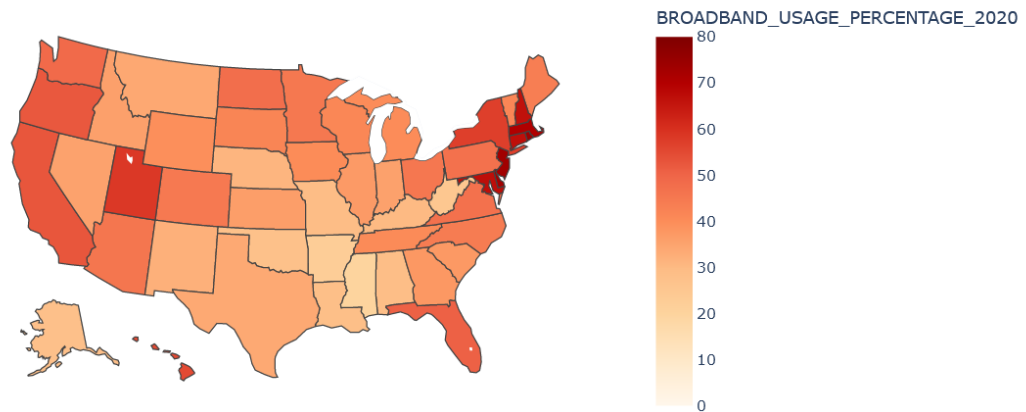


Figure 4: Map representation of Average Broadband Usage by State

Year over Year Change in Broadband Metrics

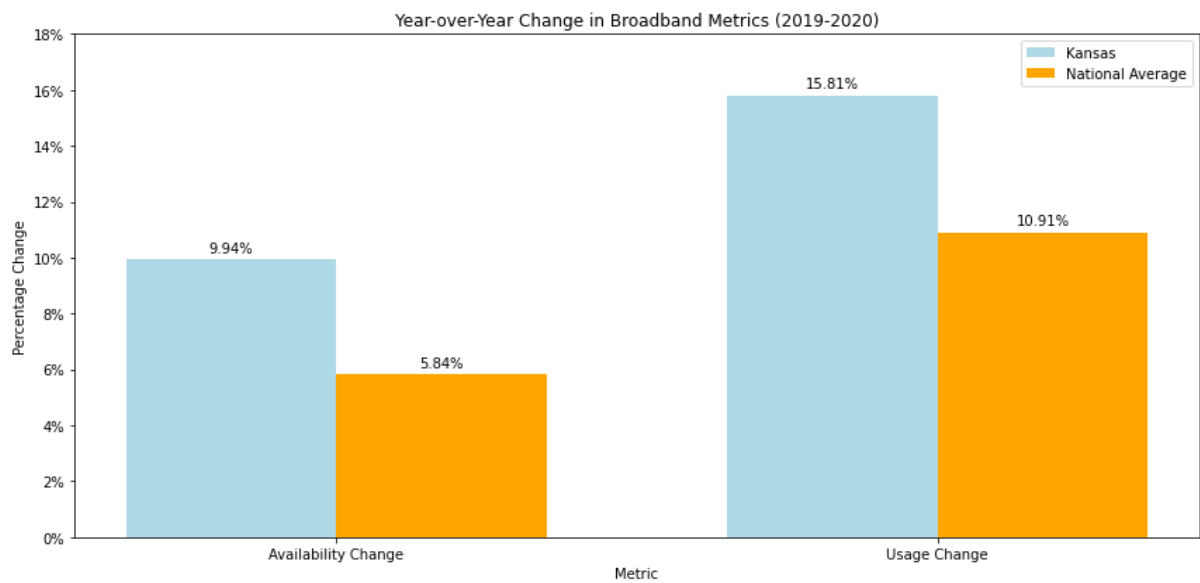


Figure 5: Year over Year Change in Broadband Metrics

- **Availability Change:** Kansas shows a significant increase in broadband availability compared to the national average, highlighting effective efforts to expand broadband infrastructure within the state.
- **Usage Change:** Similarly, Kansas's increase in broadband usage outpaces the national average, indicating growing adoption rates among residents.

This visualization demonstrates Kansas's strong performance in improving both broadband access and utilization over the analyzed period, outpacing the national average in terms of growth.

Comparison of Broadband Availability and Usage: Kansas vs National Avg (2019-2020)

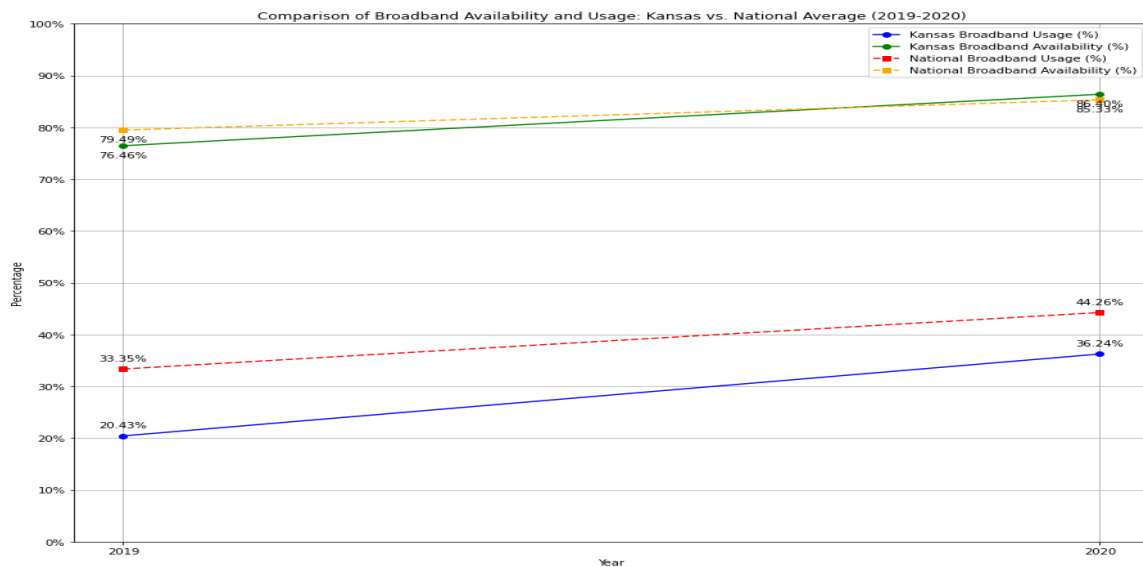


Figure 6: Comparison of Broadband Availability and Usage: Kansas vs National Avg (2019-2020)

- **Broadband Usage:** Kansas's increase in broadband usage is notably sharp compared to the national average, suggesting effective strategies or investments in digital adoption within the state.
- **Broadband Availability:** Similarly, Kansas shows a significant increase in broadband availability, outpacing the national average growth. This reflects successful initiatives to expand broadband infrastructure.

Error Metrics Evaluation

A detailed analysis of error metrics was conducted to assess the reliability of the broadband data:

- **Mean Absolute Error (MAE):** This metric helped in understanding the average error magnitude across data points, indicating how much deviation existed from true values on average.
- **95% Confidence Interval Width:** This measure provided insights into the uncertainty associated with the data points. A wider interval suggested greater uncertainty and potential issues in data accuracy.
- **Mean Squared Deviation (MSD):** This metric highlighted the variance in squared deviations, offering a view into the consistency of error magnitudes.

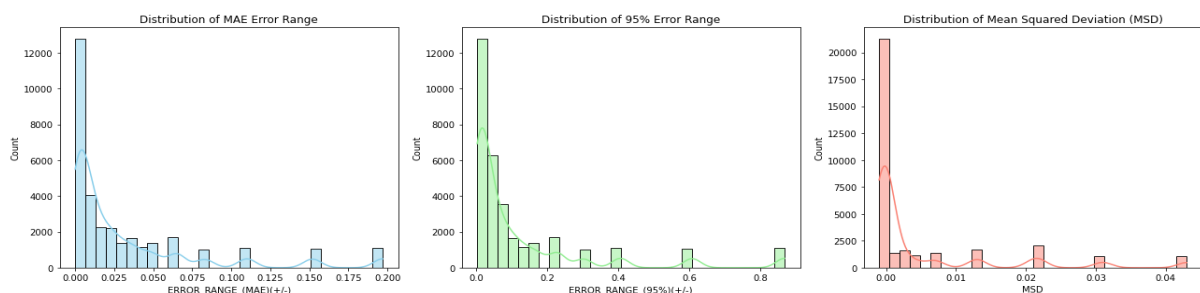


Figure 7: Distribution plots for Error Metrics

Average MAE Error Range by State

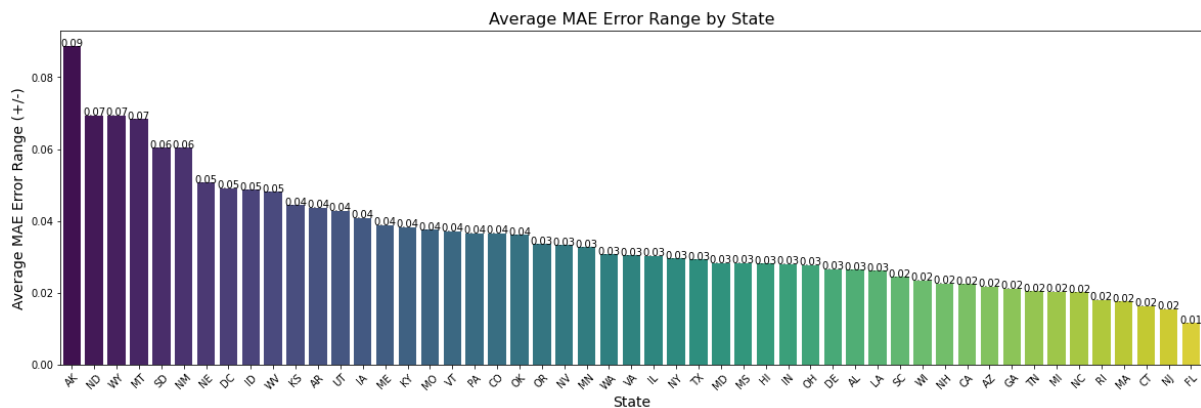


Figure 8: Average MAE Error Range by State

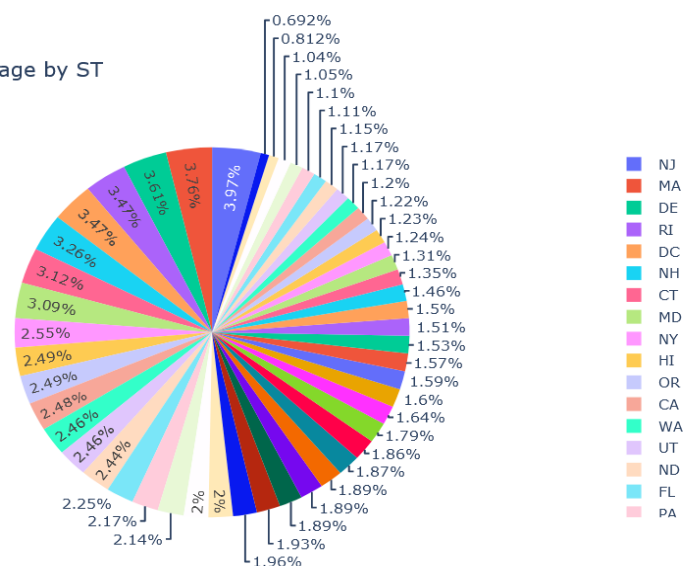
The heatmap visualizes the average Mean Absolute Error (MAE) Error Range by state, providing a clear indication of which states have higher average error measures in their broadband data. States on the left side of the heatmap, with darker colors, have higher average MAE error ranges, suggesting potential challenges in data accuracy or variability in these regions.

Insights:

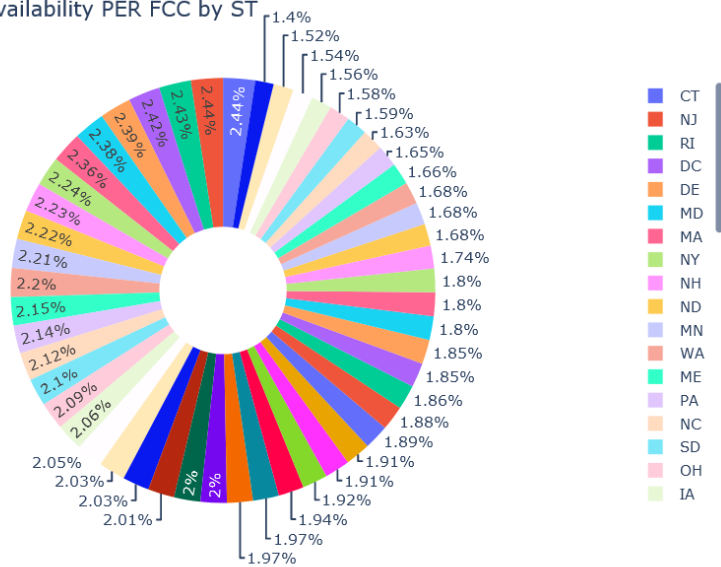
- This visualization highlights geographic variations in the reliability of broadband data across different states. States with higher average error measures may require closer examination to understand the reasons behind these errors, which could include factors like rural geography, lower population density, or challenges in data collection.
- Identifying states with higher error measures can help prioritize areas for further investigation, potential improvements in data collection methodologies, and targeted actions to enhance broadband infrastructure and access.

DASHBOARD 2019

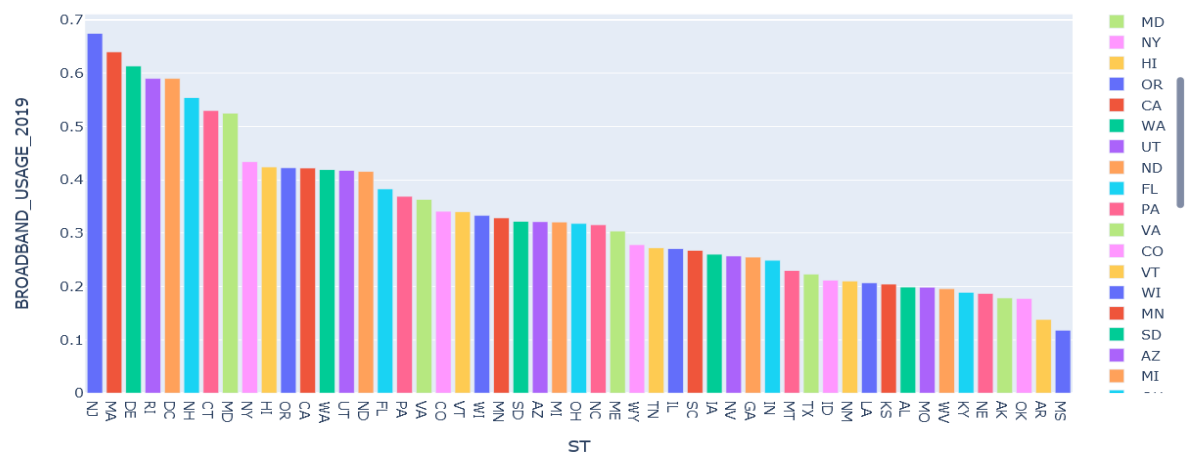
2019 Average Broadband Usage by ST



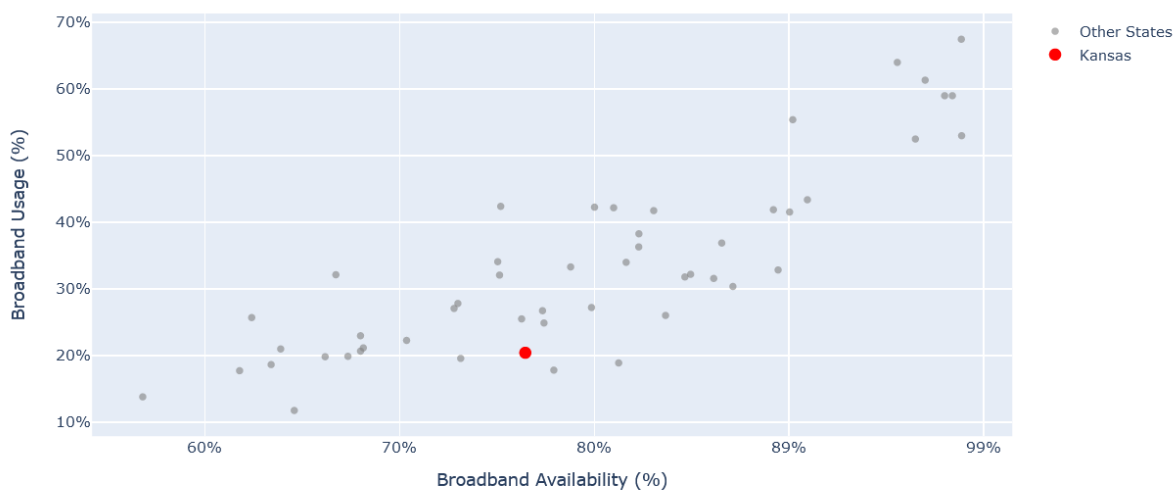
2019 Average Broadband Availability PER FCC by ST

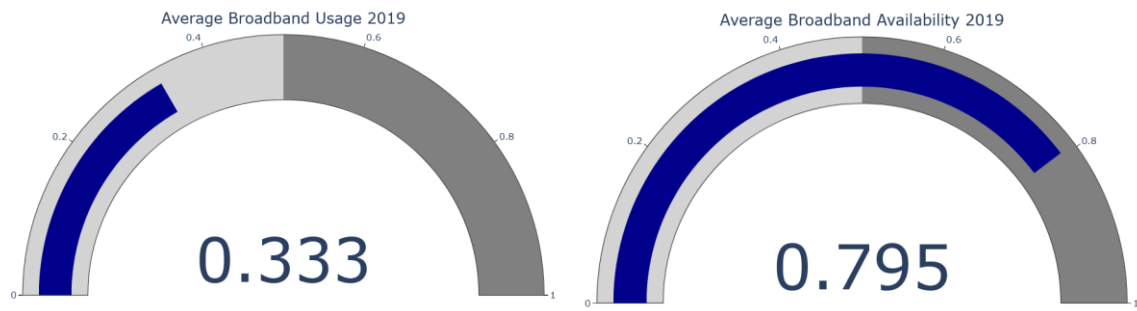


Average Broadband Usage 2019 by ST

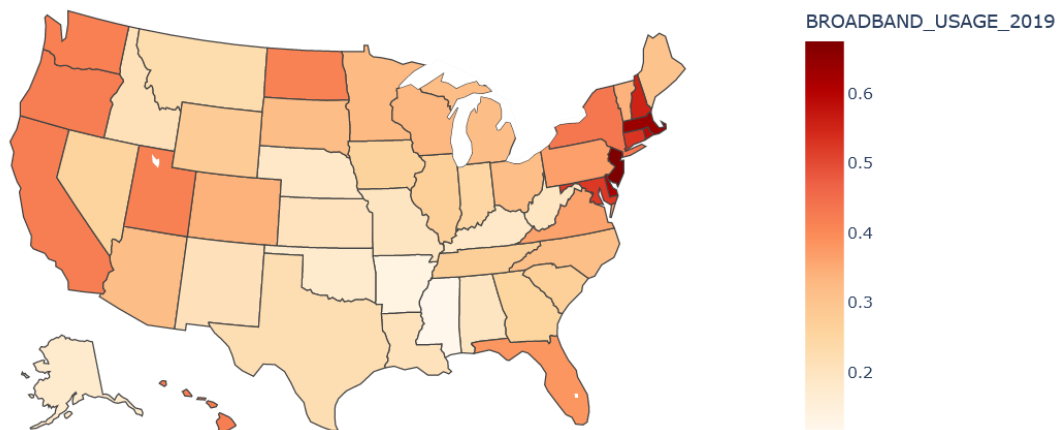


State Comparison of Broadband Availability vs. Usage (2019)



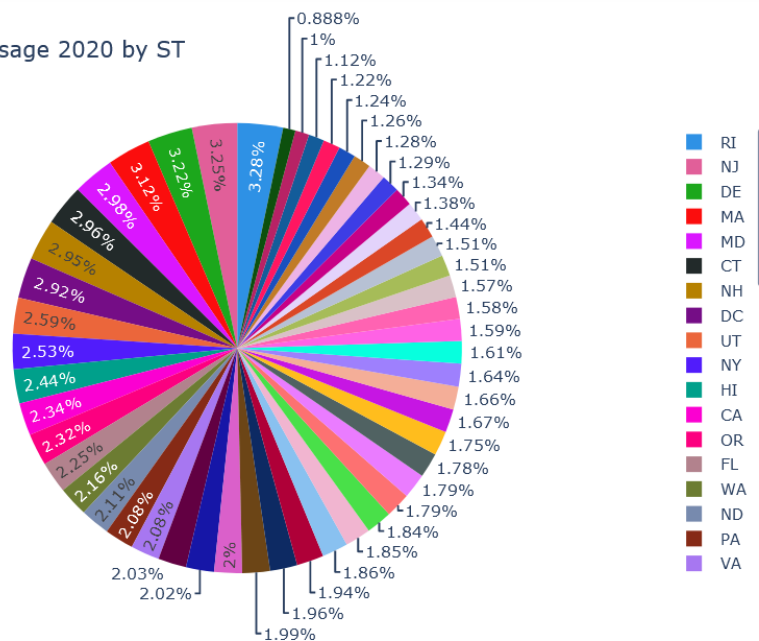


Average Broadband Usage by State (2019)

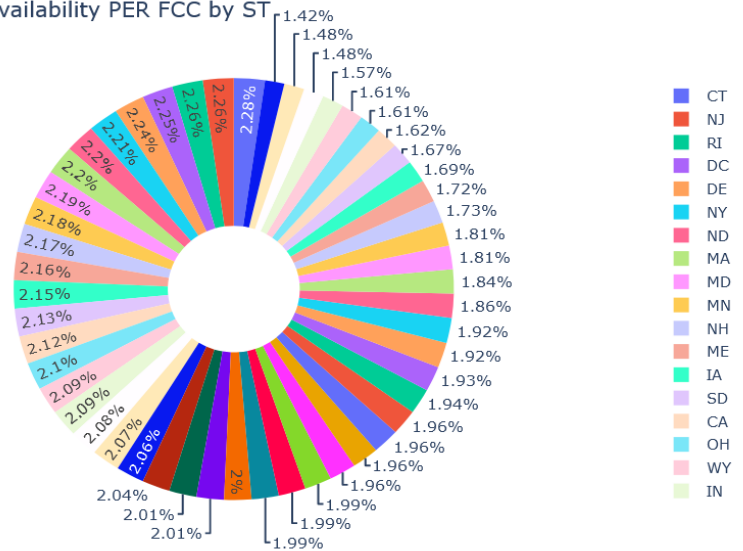


DASHBOARD 2020

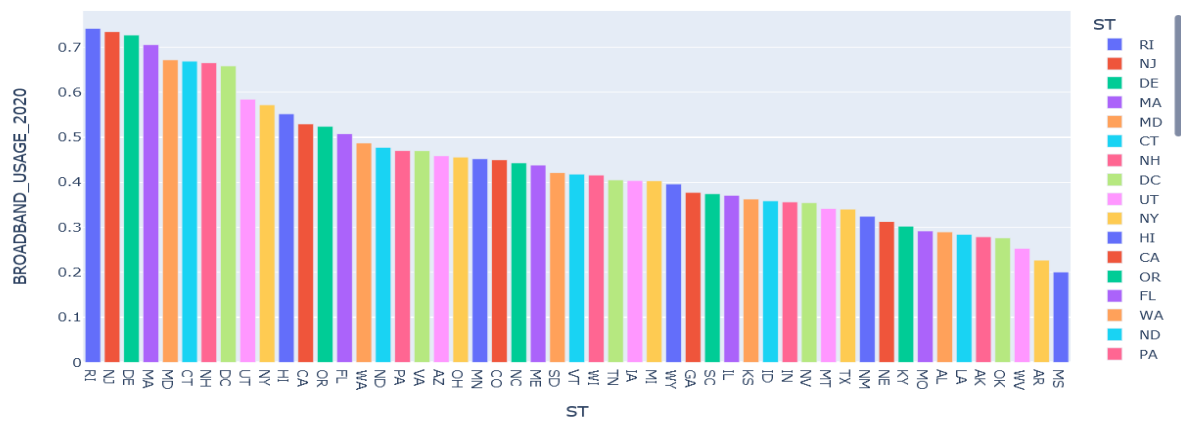
2020 Average Broadband Usage 2020 by ST



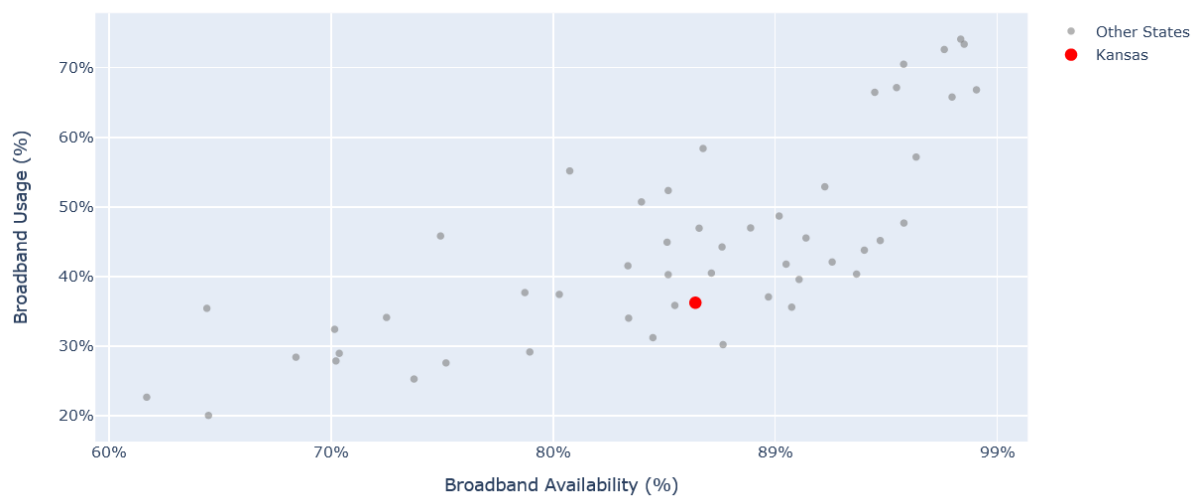
2020 Average Broadband Availability PER FCC by ST

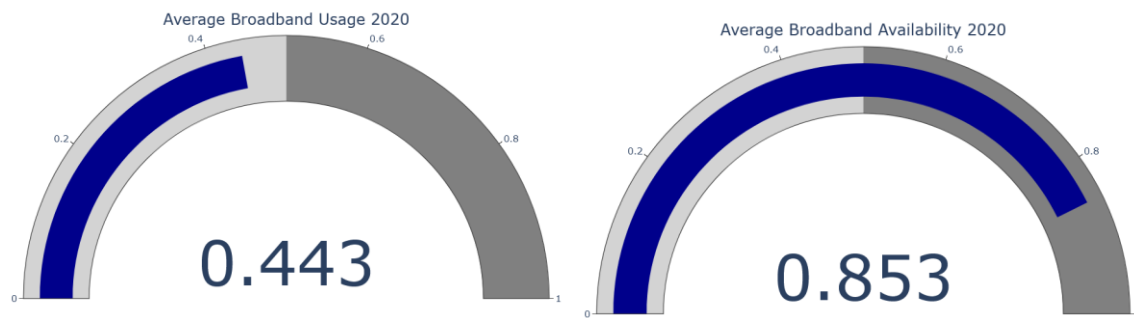


Average Broadband Usage by ST

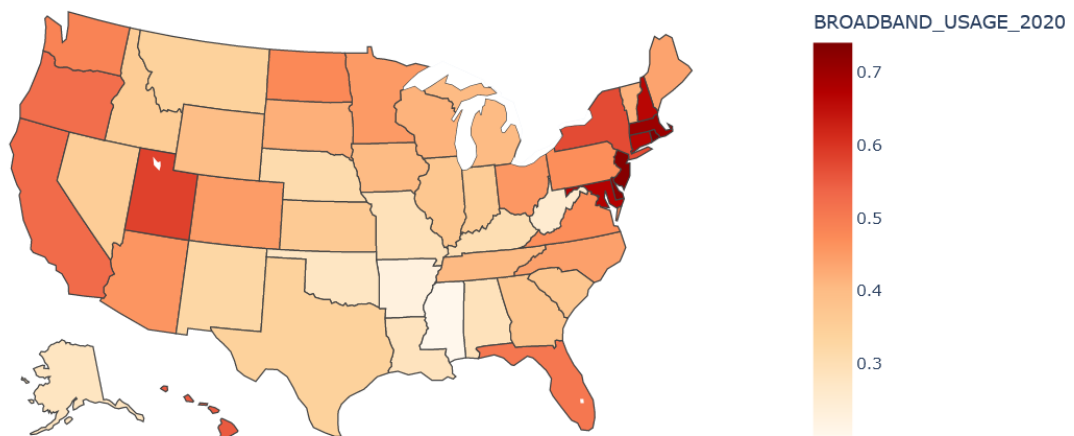


State Comparison of Broadband Availability vs. Usage (2020)





Average Broadband Usage by State (2020)



RESULTS

The results of the KBOD Analysis project provided a nuanced understanding of broadband availability and usage across different states and timeframes. Key findings from the analysis include:

Broadband Usage Increase

A notable increase in broadband usage was observed from 2019 to 2020. This uptrend is significant, reflecting either a genuine increase in broadband deployment or a shift in usage patterns possibly driven by circumstances such as the COVID-19 pandemic, which necessitated greater reliance on digital connectivity for remote work, education, and entertainment.

Geographic Disparities in Broadband Access

The analysis revealed substantial geographic disparities in broadband access:

- **Rural vs. Urban:** Rural areas consistently showed lower broadband availability compared to urban centers. This digital divide highlights the challenges in deploying broadband infrastructure in less densely populated areas.

- **State Variability:** States like Kansas and California showed different trends in broadband metrics, reflecting the impact of local policies, economic conditions, and geographic factors on broadband deployment.

Data Collection Inconsistencies

Differences in error metrics across states indicated varying levels of data collection accuracy. High error rates in certain states suggest potential inaccuracies in data reporting or collection methods, which could mislead policy decisions and infrastructure investments.

Temporal Trends and Anomalies

- **Yearly Comparison:** The comparative analysis between 2019 and 2020 data helped in identifying trends over time, showing areas with improved broadband access as well as regions where progress may be stagnating.
- **Anomalies in Data:** The error analysis highlighted specific anomalies where the confidence intervals were unusually wide, suggesting areas where broadband data might be less reliable.

Visualization Insights

Interactive visualizations provided critical insights into the data:

- **Choropleth Maps:** These maps were particularly effective in displaying broadband usage rates across counties, visually identifying areas of high and low connectivity.
- **Histograms and Scatter Plots:** These visual tools helped in understanding the distribution of broadband metrics and the relationship between availability and usage, respectively.

Implications of Error Metrics

The distributions of MAE, MSD, and confidence intervals revealed:

- **Presence of Outliers:** Certain data points with exceptionally high error measures indicated potential outliers or errors in data collection.
- **Long Tails in Distributions:** These suggest that while most of the dataset is reliable, there are significant exceptions, which could impact the overall accuracy of broadband data assessments.

RECOMMENDATION

Based on the findings of the KBOD Analysis project, several targeted recommendations are proposed to address the challenges identified in broadband availability, usage, and data accuracy. These recommendations are designed to enhance infrastructure, improve data collection methods, and ensure equitable access to broadband across all regions.

Enhance Data Collection Methods

- **Standardize Data Collection Protocols:** Implement uniform data collection standards across states to ensure consistency and reliability of broadband data. This would help in reducing the discrepancies and high error rates observed in the data analysis.
- **Regular Data Audits:** Conduct regular audits to verify the accuracy of the data collected and address any anomalies or errors promptly. This will help maintain the integrity of the data used for analysis and decision-making.

Target Infrastructure Improvements

- **Prioritize Rural and Underserved Areas:** Focus infrastructure improvements on rural and underserved areas where broadband availability is notably low. Investments could include expanding fiber-optic networks, supporting satellite broadband solutions, and incentivizing private sector involvement.
- **Leverage Public-Private Partnerships:** Encourage partnerships between government bodies and private companies to pool resources and expertise in deploying broadband infrastructure more efficiently.

Implement Policy Enhancements

- **Develop Incentive Programs:** Create incentive programs for broadband providers to expand services in underserved areas. These could include tax breaks, subsidies, or grants that make it economically viable to build infrastructure in low-density areas.
- **Update Regulatory Frameworks:** Modernize regulatory frameworks to facilitate faster and more extensive broadband deployment. This includes easing right-of-way restrictions and supporting the deployment of new technologies such as 5G.

Foster Community Engagement and Awareness

- **Community Outreach Programs:** Initiate community outreach programs to raise awareness about the benefits of broadband and digital literacy. These programs can also gather community feedback which can guide infrastructure development based on local needs.
- **Educational Initiatives:** Work with educational institutions to provide programs that enhance digital skills, preparing communities to better utilize broadband services for education, healthcare, and business.

Monitor Progress and Adjust Strategies

- **Implement Monitoring Systems:** Set up systems to continuously monitor broadband deployment and usage. This real-time data can help in quickly identifying areas where strategies may need adjustment.
- **Flexible Policy Adjustments:** Maintain a flexible approach to policy-making that allows for quick adjustments based on new data and evolving technological landscapes.

Conclusion

The KBOD Analysis project has provided critical insights into the state of broadband availability and usage across various regions in 2019 and 2020. Through comprehensive data analysis, significant disparities in broadband access have been highlighted, particularly between urban and rural areas, alongside notable fluctuations in data accuracy. The project's findings underscore the complexity of broadband deployment and the urgent need for strategic interventions to address both infrastructural deficits and data integrity issues.

Summary of Key Findings

- **Increased Broadband Usage:** The analysis revealed an upward trend in broadband usage, reflecting greater societal reliance on digital connectivity. However, this increase is not uniformly distributed, with rural areas lagging significantly behind urban centers.
- **Geographic Disparities:** The project identified stark geographic disparities in broadband access, which pose a substantial barrier to achieving equitable digital inclusion.
- **Data Collection Challenges:** Variability in error metrics across different states highlighted the challenges in data collection, emphasizing the need for standardized, accurate data to inform policy and investment decisions.

Implications for Stakeholders

The implications of these findings are profound for various stakeholders, including policymakers, telecommunications providers, community leaders, and educators. For policymakers, the project outlines the critical areas needing legislative attention and resource allocation to bridge the digital divide. Telecommunications providers are called upon to expand their infrastructure, particularly in underserved areas, potentially spurred by incentive programs. Community leaders and educators can leverage these insights to advocate for better resources and to tailor educational programs that enhance digital literacy and utilization.

Path Forward

Moving forward, the recommendations provided aim to catalyze actions that will ensure more equitable broadband access. This involves not only significant infrastructural investments but also a concerted effort to improve the quality and accuracy of broadband data collection. The proposed enhancements in data collection methods and infrastructure deployment strategies are expected to foster a more connected, inclusive digital environment.

Final Reflections

In conclusion, the KBOD Analysis project highlights the vital role of broadband in today's society and the disparities that persist in its access. It calls for a unified approach involving

various stakeholders to implement effective solutions that address both the infrastructural and data-related challenges. By embracing these recommendations, there is a promising path forward towards a more digitally inclusive society, where every individual can benefit from the opportunities provided by broadband connectivity.

The learnings from this project should serve as a blueprint for ongoing and future efforts to enhance broadband deployment and utilization, ensuring that broadband as a utility is as ubiquitous and essential as electricity and water, and just as accessible.

References

1. McKinney, Wes. 2017. "Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython." 2nd ed. Sebastopol: O'Reilly Media.
2. Hunter, John D. 2007. "Matplotlib: A 2D Graphics Environment." *Computing in Science & Engineering* 9 (3).
3. Waskom, Michael. 2021. "Seaborn: Statistical Data Visualization." *Journal of Open Source Software* 6 (60).
4. Caine, Carolyn. "Research on broadband in Kansas recommends statewide strategy to improve infrastructure." University of Kansas News, March 22, 2023
5. Oxford Academic. "The Impact of Broadband Access on Social and Economic Outcomes." *Journal of Refugee Studies*, vol. 34, no. 1, 2021