

# NATIONAL BROADBAND TOOL

Empowering Policymakers to Close the Digital Divide

Shweta Sampath Kumar, Swathi Ganesan, Kishor Mannur, Marc Edwards Mentored by: Donald Patchell



O1
INTRODUCTION

O2
BUSINESS PROBLEM

O3

DASHBOARD

O4

DATA OVERVIEW

O5
METHODOLOGY

OG OUTCOMES

O7
IMPACT

O8

FUTURE WORK

Q&A



# **Broadband's Importance in Society**



**Boosts Economic Growth** 



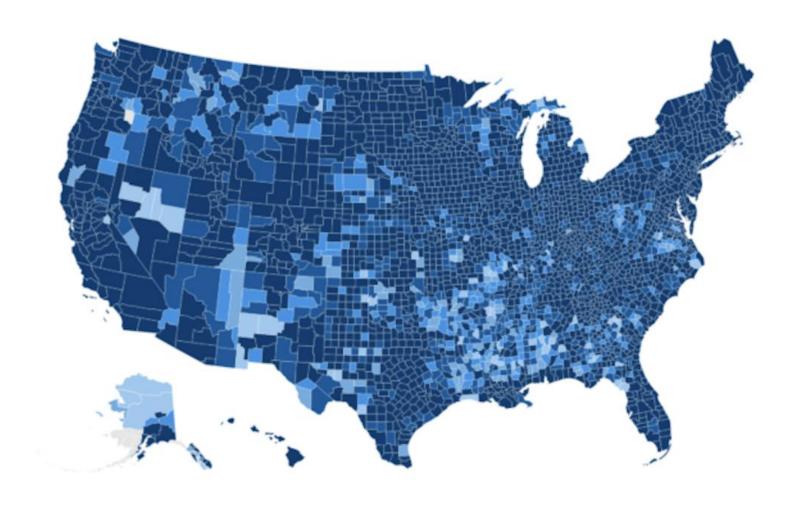
**Enhances Educational Access** 



**Supports Healthcare Advancements** 



**Expands Employment Opportunities** 





<sup>\*</sup> FCC Broadband has or "could" provide greater than or equal to 25 Mbps / 3 Mbps

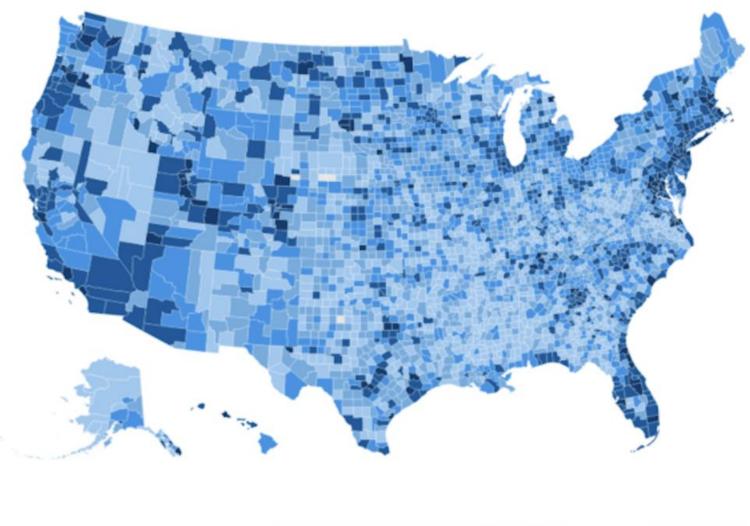
# Current Landscape of High-Speed Broadband Access in the USA

According to the Federal Communications Commission (FCC),

~14.5M

people <u>DON'T</u> have access to internet at broadband speeds

INTRODUCTION BUSINESS PROBLEM DASHBOARD DATA OVERVIEW METHODOLOGY OUTCOMES





<sup>\*\*</sup> Broadband speeds greater than or equal to 25 Mbps

# Current Landscape of High-Speed Broadband Access in the USA

.....but according to Microsoft,

~120.4M

people <u>DON'T</u> have access to internet at broadband speeds

# Nationwide Broadband: A Strategic Leap Forward

Infrastructure Investment and Jobs Act -The Bipartisan Infrastructure Deal

### **NOVEMBER 15, 2021**

~\$65B fund into high-speed internet efforts through 7 federal programs aiming to modernize and upgrade the country's infrastructure and ensure everyone in America has access to Internet service



Biden-Harris Administration Announces Over \$42 Billion to Connect Everyone in America to Affordable, Reliable, High-Speed Internet

#### **JUNE 26, 2023**

- Varied awards in the \$27M to \$3.3B range
- 19 states exceed \$1B allocations
- Goal: Universal Access by 2030



# **US Ignite: Bridging the Digital Divide**



A non-profit organization founded in 2012 based in Washington DC with a mission

To forge private-public partnerships that accelerate the adoption of advanced technologies, catalyzing smart infrastructure and socio-economic growth in the United States

# Addressing the Rural Broadband Challenge







### Challenge

Identifying potential
broadband infrastructure
investment opportunities in
broadband deserts with
promising socio-economic
growth trajectories



A National Broadband Tool with interactive maps and charts to visualize and analyze the transformative impact of high-speed internet access

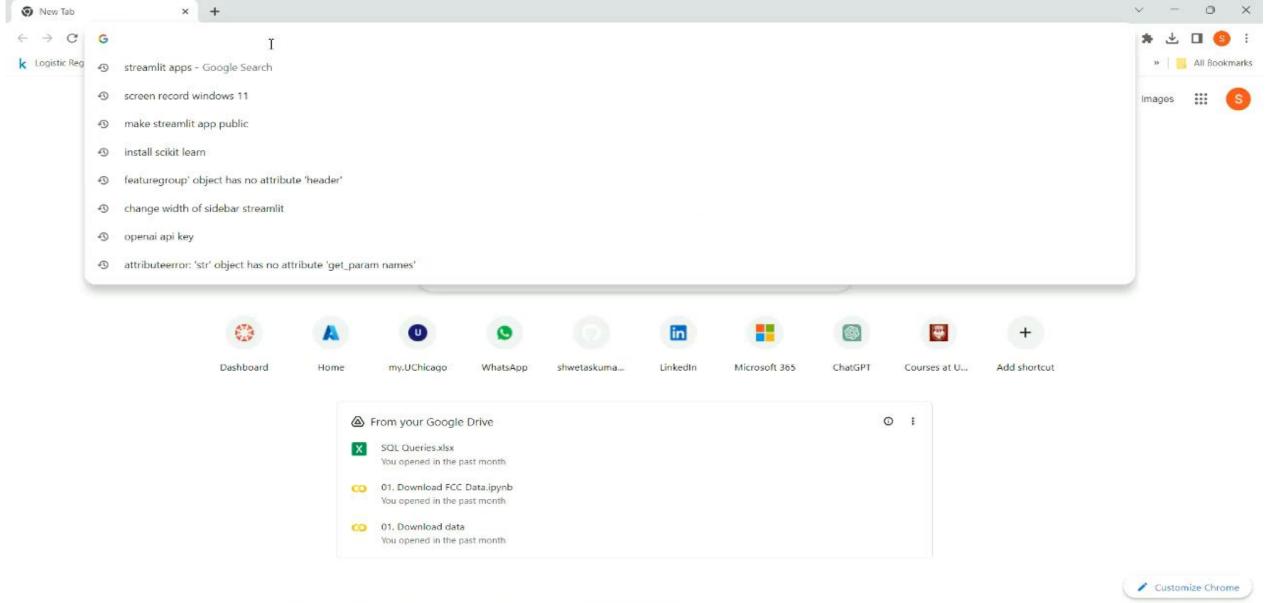
#### Outcome

Assisting community leaders in developing a compelling proposal to secure grant funding by effectively highlighting the benefits of investing in broadband infrastructure



**INTRODUCTION BUSINESS PROBLEM DASHBOARD** DATA OVERVIEW **METHODOLOGY** OUTCOMES

### **DASHBOARD DEMO**





























### **DATA SUMMARY**





2017 - 2021

50 states

16 territories

84,414 census tracts

#### Census Bureau

#### CENSUS BEREAU – AMERICAN COMMUNITY SURVEY

#### Demographics

- Gender
- Age
- Veteran Status
- Disability Status

#### **Education**

- High School
- Bachelors
- Masters
- No degree obtained

#### **Employment Status**

- Employed
- Not Employed

#### **Median Income**

- Less than \$20k
- \$20k \$75k
- \$75k+

#### **Broadband Access**

- Internet subscriptions by households
- Type of connections:
  - Broadband
  - Other + satellite
  - No Internet



### FEDERAL COMMUNICATIONS COMMISSION (FCC)

#### **Internet Provider Filings**

- Connection Type: Broadband, Satellite, Other
- Number of connections with download speed at least 25 Mbps
- Number of connections with download speed less than 25 Mbps

#### **Area Shape Files**

Geospatial Census Data Representation

### **DATA HANDLING**

#### DATA CHALLENGES

Using vast volumes of data sourced from open-access datasets submitted by Internet Service Providers (ISPs) to the FCC every six months, combined with ACS Census Data to support broadband policy & investment decision-making and helping policymakers direct funding and programmatic investments to communities most impacted by the digital divide.

**Exclusion Criteria** 

Omitted non-residential census sectors such as recreational zones, commercial areas, and woodlands.

**Handling Missing Data** 

Substituted absent data points with zero to align with the standardized format of the data presentation on the portal.

**Improved Column Labels** 

Updated to reflect their true definitions, originally labeled with Census Data codes such as **B01003\_001E** = "**Total Population**".

### **DATA PREPROCESSING**

#### **Broadband Connections**

 $Digital\ Index = \frac{\#\ broadband\ subscriptions \times \#\ connections\ at\ least\ 25mbps}{total\ \#\ of\ connections}\ \div\ total\ broadband\ connections\ available$ 

### Population, Income, Employment and Education Statistics

Captured demographic details as a **percentage of total population**, encompassing **gender**, **age**, **education**, **and income brackets**.

#### **Rural Areas**

An area is tagged as rural\* if it contains less than 2,000 housing units & has population less than 5,000.

#### **Shape Files**

Converted shape files to **GeoJSON** format for building maps.

BUSINESS PROBLEM DASHBOARD DATA OVERVIEW METHODOLOGY OUTCOMES IMPACT

### MODELLING I – PREDICTING SOCIOECONOMIC GROWTH

### AIM

To determine potential growth in socio-economic status (across target variables) on account of increased broadband access

#### MODEL

**XGBoost Regression With Random Search using Cross Validation** 

#### **INPUT PARAMETERS**

- Demographics Gender, Age, Veteran Status, Disability
- 2 Yr. Lagged Dependent Variables for Causality Education, Employment, Income
- % of broadband connections at least 25 mbps
- % of broadband connections less than 25 mbps
- % type of connection **Broadband, Other, No internet**

#### TARGET VARIABLES



### MODELLING II – IDENTIFYING DEMOGRAPHIC CLUSTERS

### AIM

Identifying demographically similar census tracts to help make well-informed decisions regarding investments in broadband infrastructure

#### MODEL

**K-Means Clustering** 

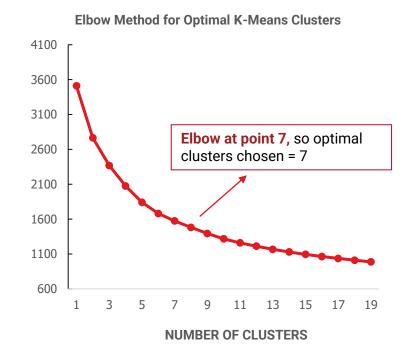
STATE: ALABAMA | COUNTY: Marengo | TRACT - 01091973400 | Digital Index - 0.54

Population	Male	Female	White	Black	Asian	Native	Pacific	Other	Multi- Race	Under 25yrs	Over 25yrs	Seniors	Veterans	Non- Veterans	ctatus	Disabled status above 65 yrs	
3688	50.41	49.59	73.21	20.44	2.79	0.24	0	0	3.31	22.97	48.02	29.01	11.5	71.12	16.7	16.92	

#### **OUTPUT**

tract_code	county	Population	Male	Female	
04013612700	Maricopa County	2273	52.05	47.95	
19079960500	Hamilton County	1939	47.5	52.5	
27111960500	Otter Tail County	5668	47.64	52.36	
35001000112	Bernalillo County	2040	49.31	50.69	
38067950500	Pembina County	1390	51.8	48.2	

Veterans	Non- Veterans	Disabled status below 65 yrs	Disabled status above 65 yrs	Digital Index
7.74	76.2	4.75	6.03	1
6.29	73.23	7.99	4.13	1
5.42	70.43	7.2	8.27	1
10.34	71.81	4.66	6.27	1
6.83	72.66	5.04	5.32	1
				\ /



### **MODELLING III – NATIONAL BROADBAND GPT**

### AIM

To assist policymakers in understanding the statistics of their census tract & create grant proposals for broadband infrastructure grants.

#### TRAINED DATA

#### Sample Grant Proposal

Use this sample grant proposal to help you in the process of preparing your own request for funding. Check with your state department of education to see if they require a certain format in their request for proposal (RFP). Many foundations also have specific requirements.

This sample grant proposal contains:

- Sample Cover Letter
- Sample Cover Page
- Sample Grant Proposal
- Sample Letter Format Grant Proposal for Foundations

# Knowledge Conversations with your GPT may include file contents. Files can be downloaded when code interpreter is enabled. grant\_template.pdf PDF Bipartisan deal\_info.pdf PDF The following files are only available for Code Interpreter: visual\_data.csv Spreadsheet

#### **SAMPLE PROMPTS**

Hi, please help me design a grant proposal for my tract {tract code}. I am a public official, could you help in retrieving the required information to build a grant proposal and build a draft proposal.

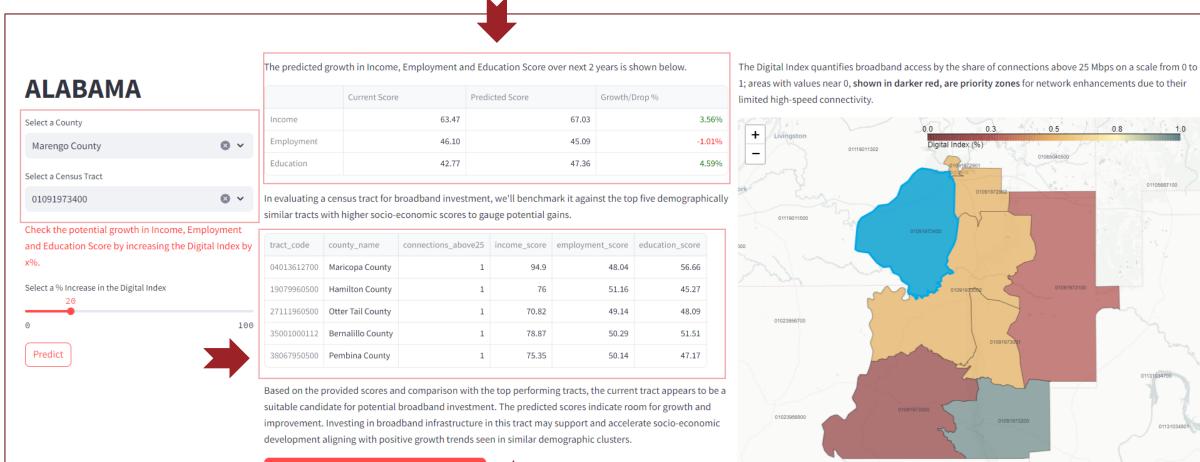
Considering the number of broadband connections, do we satisfy the requirements for the act?

Could you provide more information on this tract relevant to applying for highspeed broadband access?

Whom should we approach for further steps to help improve our tract's digital index?

**BUSINESS PROBLEM DASHBOARD METHODOLOGY OUTCOMES DATA OVERVIEW IMPACT** 

### Our Solution in Action - National Broadband Tool



Chat with NBT.AI to build your grant proposal!

9/0		0.0	0.3	0.5	0.8	1.0
%	+ Livingston 01119011302	Digital Inde	x (%)		~ '. K.	
%	-	01091	972901	01085040500	- 100	
	Drik		01091972902			01105687100
ally	The state of the s		7			
	01119011500	01091973400	7			
	300	01031313400				
	Jan 1 money				1, 1, 1	
		01001	sapske	01091973	. 2	
		J. J	- Janes		1	
	01023956700					
			0109197	3001		Same
		No. of the last of	6			
			-		011	31034700
e a						7 1
С	01023956800	01091973300	0109	1973200		01131034801
		200				
					1131035100	
	00			0	1131033100	
			Leaflet   ©	OpenStreetMap contri	outors © CartoDB, Car	toDB attributions



Census Tract	County	Population	Male	Female	White	Black	Asian	Native	Pacific	Other	25+ yrs old without graduating high school	Without an internet subscription, broadband of any type	Households without a desktop or laptop	Digital Index	% of people not using internet at b
01091973400	Marengo County	3,688	50.41	49.59	73.21	20.44	2.79	0.24	0	0	5.17	14.39	11.12	0.54	

BUSINESS PROBLEM DASHBOARD DATA OVERVIEW METHODOLOGY OUTCOMES IMPACT

# Our Solution in Action – LLM Generated Proposal Inputs



#### **National Broadband GPT**

This GPT helps public officials in understanding the statistics of their tract code & create grant proposals for broadband infrastructure act. Find more Information here:

https://www.whitehouse.gov/briefingroom/statements-releases/2021/11/06/fact-sheetthe-bipartisan-infrastructure-deal/

Hi, please help me design a grant proposal for my trac...

Considering the number of broadband connections, d...



Message National Broadband GPT...



Let's proceed to create a draft of the grant proposal. hinspace

#### **Grant Proposal Draft for Tract 01091973400**

#### Project Abstract

The "Community Advancement Initiative" aims to significantly improve the quality of life in Tract 01091973400 through targeted interventions in income, employment, and education sectors. The current and predicted data for these areas indicates a pressing need for strategic investments and programs. This initiative seeks to bridge gaps, foster growth, and create sustainable development in these key domains.

#### Statement of Need

The tract has shown mixed trends in the key areas of income, employment, and education. While there's a predicted increase in income (3.56%) and education (4.59%), employment is expected to decline (-1.01%). These statistics underscore the need for a comprehensive strategy to bolster economic stability and educational opportunities while countering the downward trend in employment.

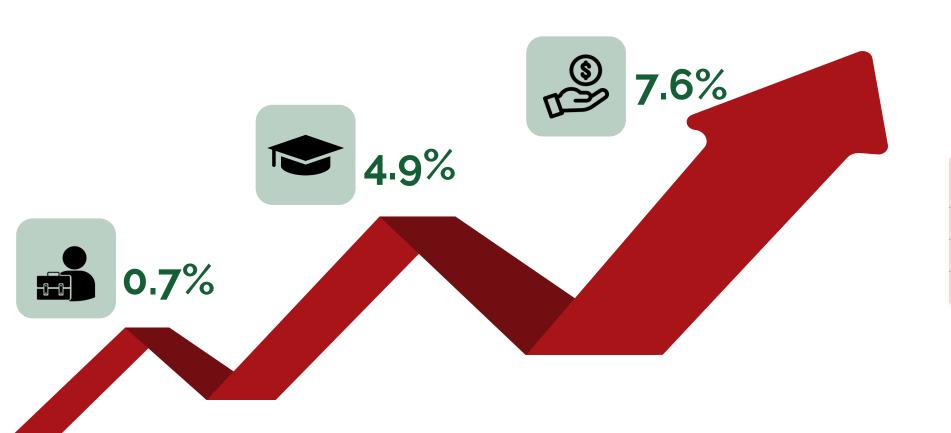
#### **Program Description**

The program will consist of three primary components:

- 1. **Economic Empowerment**: Initiatives to increase income levels, such as small business support, skill development programs, and investment in local enterprises.
- 2. **Employment Opportunities**: Addressing the predicted drop in employment through job fairs, vocational training, and partnerships with local businesses for job creation.

### Bridging Digital Equity Gaps in Underserved Rural Regions via Broadband

A **20% increase** in **25 Mbps broadband connections** in **rural** areas holds the potential to significantly elevate **employment**, **education**, **and income** within a span of **two years** 



#### **Accuracy and Reliability**

TARGET	RMSE
EDUCATION	0.066538
EMPLOYMENT	0.015154
INCOME	0.036163

### **Future Work**

# Collating Enhanced Broadband Data

Refine identification of Broadband Serviceable Locations (BSL) by distinguishing between served, underserved, and unserved areas using data from additional partner states, federal agencies, industry, and accessible commercial datasets

# **Community Report Generation**

Automate end to end grant proposal generation by creating community reports offering proposed broadband and socioeconomic metrics to bridge digital equity gaps

#### **Sustainability Planning**

Explore P3 models and funding strategies to ensure the long-term sustainability of the tool, including potential partnerships with broadband providers

# The Team







MARC



**SWATHI** 



**KISHOR** 



DON PATCHELL ADVISOR

**OUR CLIENT PARTNER** 



# **THANK YOU**

# **APPENDIX**

# **APPENDIX – US Ignite**

#### **BUSINESSES**



Need repeatable, scalable, and sustainable commercial strategies in the emerging smart community market.

#### **LOCAL GOVERNMENTS**



Must enhance quality of life and drive economic development amid rapid technological change.

#### **FOUNDATIONS & FEDERAL AGENCIES**



Align institutional aims with research and economic development initiatives that deliver measurable benefits.

#### **INDUSTRY PARTNERS**





A non-profit organization founded in 2012 to accelerate the adoption of advanced technologies in the United States



#### UNIVERSITIES





# **APPENDIX – US Ignite**

#### **Approach**

#### **Public-Private Partnerships**

Collaborating with communities, universities, and industry partners to foster a smart city ecosystem.

#### **Expertise & Innovation**

A team with deep technical expertise and a track record of innovation across the smart city spectrum.

### **Community Benefits**

#### **Ecosystem Integration**

Join an ecosystem of connected communities, innovators, and transformative applications.

#### **Support & Development**

Access technical assistance and community-building support, including funding opportunities.

### **Mission & Impact**

#### **Strategic Mission**

To build impactful applications requiring gigabit fiber networks and to incubate these within test bed communities.

#### **Smart City Development**

Facilitating public-private collaborations for smart city progress and economic growth.

#### **Diverse Partnerships**

Corporate, government, and foundation partners united for smart city initiatives

**Clients and Partners** 

#### **National Effort**

Over **40 cities** engaged in driving the smart city movement



# **APPENDIX** -Background

### **DIGITAL DIVIDE**

According to the Federal Communications Commission (FCC)

**21.3** M

people don't have internet access

**40%** of schools

60% of health care facilities

Outside metropolitan areas lack broadband

### **Infrastructure Investment and Jobs Act**

The *Infrastructure Investment and Jobs*Act (effective November 15, 2021) is a bipartisan bill signed into law that aims to modernize and upgrade the country's infrastructure.

\$65 BILLION

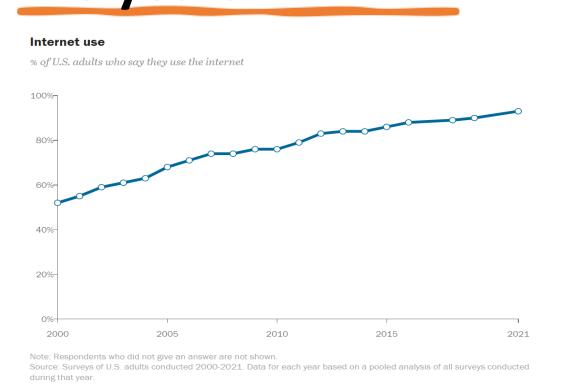
Broadband deployment

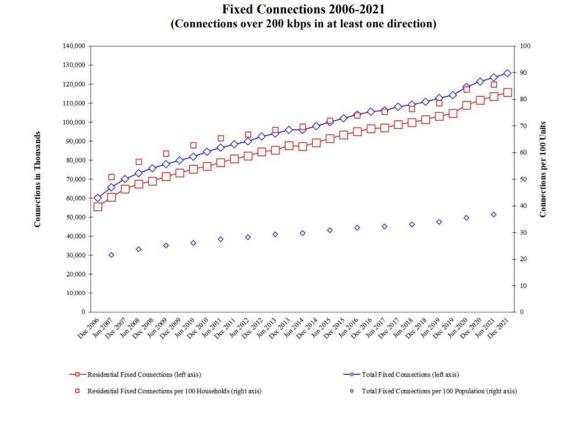


Through its programs, US Ignite can leverage the funding provided to help connect communities and bridge the digital divide.



# **APPENDIX** – Change in broadband access over the years





These graphs illustrate the significant growth in internet usage among U.S. adults from 2000 to 2021, revealing a near-saturation point as the percentage approaches full coverage by 2021. It highlights the pivotal shift towards a digitally connected adult population over two decades.



### APPENDIX - MODELLING III - NATIONAL BROADBAND GPT

AIM

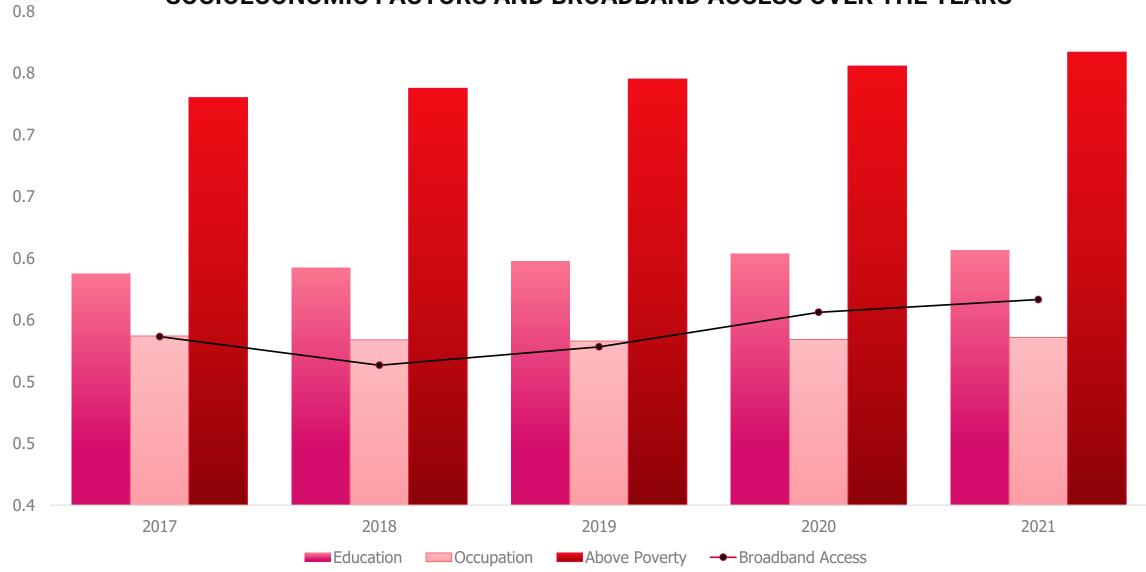
This GPT helps public officials in understanding the statistics of their tract code & create grant proposals for broadband infrastructure act.

#### INSTRUCTIONS

You are a GPT that is designed to help public officials and homeowners petition the government via grant proposals for the Broadband Provisions of the 2021 Infrastructure Investment and Jobs Act. You must act in a formal and guided approach and should not respond to irrelevant or inappropriate messages. For data-specific inquiries, you should use the provided datasets and scrape the below mentioned sites for information. When creating grant proposals, use the given grant proposal document to format the template.

### **APPENDIX – Exploratory Data Analysis**

#### SOCIOECONOMIC FACTORS AND BROADBAND ACCESS OVER THE YEARS



# **Appendix - Data**

### **Candidates for broadband data:**



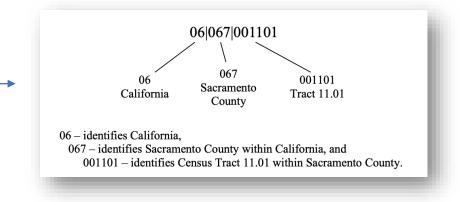
- Form 477 Data | Federal Communications Commission
- National Telecommunications and Information Administration

#### 

### Demographic, geographic and economic index:



- Census Bureau Data | American Community Survey
- Bureau of Labor statistics



**DATA SOURCES:** All datasets that will be analyzed will be sourced from open-source datasets.



# **Appendix - Project Plan**













#### **VALIDATION**

- Validate the clusters previously designed
- Gain a deeper understanding of their characteristics through statistical tests and visualizations

#### **NEW ERA**

- Collect new data if available
- Re-run and fine tune Models
- Segment map tracts from old to new clusters to identify changes in broadband demand over time

#### **IDENTIFY ROOT CAUSE**

- Identify key factors that have
   driven changes in broadband
   demand over time
- Find stories by identifying
   significant correlations or patterns

#### **GET THAT GRANT!**

- Develop a compelling story and proposal for grant funding using insights gained
- **Highlight the benefits** of investing in broadband infrastructure.



# **Appendix – Project Approach**

01

02

03

04

**Identifying Broadband Deserts** 

**Visualising Peer Counties** 

**Interpreting Geospatial Clusters** 

**Data-Driven Community Report** 

Geographical clustering and predictive modelling to identify and analyze towns that require broadband access

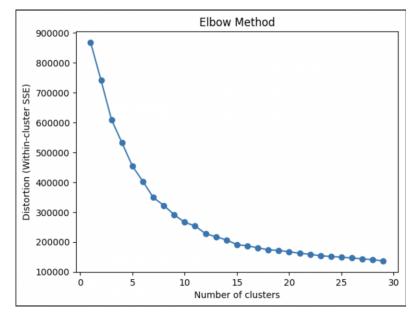
Designing a dashboard that can assist local officials & agencies in making a case to obtain funding grants

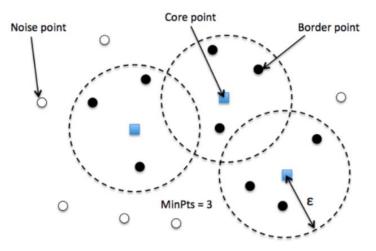
Using the dashboard, policymakers can overlay and analyse multiple variables simultaneously on an interactive map, with periodic automatic data updates

reports offering
broadband and socioeconomic metrics at a state and county level for the area of interest

## **Appendix - Identifying Demographic Clusters by US Ignite**

- We cluster the tracts into 7 clusters based on its demographic feature values for the year 2019
- Method used to arrive at the ideal number of Clusters is Elbow Method
- Algorithms experimented with for Clustering:
  - Cluster DBSCAN
  - K Means
- We used K Means because it better handles the outliers in the data
- Nearest Neighbors Algorithm to identify and store the top n-most similar tracts for a given tract based on a set of filters like State, County and Broadband Score





## **APPENDIX - Predicting Socioeconomic Growth by US Ignite**

### **Input Parameters for training**

- Socioeconomic Features
- Demographic Features/Cluster Labels
- Broadband Score
- New Broadband Score

### **Output Parameters**

New Socioeconomic Scores

### **Scaling of Input Data**

- Min Max Scaling
- Standard Scaling

### Introducing lags

 Changes in socio-economic features are expected to be seen at least 2 years after increasing the Broadband score

#### **Model Selection**

- Linear Regression
- Random Forest

ML Algo	Scaling Algo	Input Granularity	Lagged Model	Error
Linear Regression	Min-Max	Demographic Values	N	0.02691
Random Forest	Min-Max	Demographic Values	N	0.02207
Random Forest	Standard	Demographic Values	N	0.3145
Random Forest	Min-Max	Demographic Values	Y	0.0181
Random Forest	Min-Max	Clustered Values	Y	0.02068

# **Appendix - CARTO – Model Outputs**

#### 'tracts' table



tractcode number	countyname string	statename string	geoid string	cluster_label
17031071400	Cook	Illinois	14000US17031071400	11
17031071500	Cook	Illinois	14000US17031071500	11
17031071600	Cook	Illinois	14000US17031071600	2
17031071700	Cook	Illinois	14000US17031071700	2
17031071800	Cook	Illinois	14000US17031071800	2
17031080100	Cook	Illinois	14000US17031080100	8
17031080201	Cook	Illinois	14000US17031080201	8
17031080202	Cook	Illinois	14000US17031080202	8
17031080300	Cook	Illinois	14000US17031080300	2
17031080400	Cook	Illinois	14000US17031080400	6

# **Appendix - CARTO – Data Tables**

# CARTO

### 'tracts\_observations\_predicted' table

geoid string	broadband number	education_level	school_enrolled_perce number	cs_math_occupations	income number	ratio_income_to_poverty:	travel_time_to_work number	overall_socioeconomic inumber	education number	:
14000US01001020802	4	2.83498261567264	26.5873384772706	23.1454005934718	2.80881191470157	3.76551968783256	2.47748812583107	2.64601865114459	2.44923807738173	
14000US01001020900	2	2.70349519117992	22.6411903956713	9.86547085201794	2.82804757185332	4.34265734265734	2.16356675046965	2.55633888434463	2.30457140350339	
14000US01003010701	4	3.43170651817884	24.7090631244857	28	3.07869334079463	4.72747252747253	2.39358622768687	2.95663685651872	2.71003452157913	
14000US01003010905	2	2.67103054716348	22.8770359318662	0	2.74237440191388	3.64393179538616	2.51622113425316	2.41477321933189	2.29305599221906	
14000US01003011401	2	2.80951278787177	18.5404511016285	3.28947368421053	2.88671755725191	3.62065589818894	2.2256246001394	2.40261797248093	2.27556541596846	
14000US01003011403	3	3.09270307480496	19.3816125501642	0	2.87032650118447	4.09953161592506	2.63455773091061	2.57873057080528	2.43398378840576	
14000US01003011502	2	2.6441998130647	18.7344415000401	38.9067524115756	2.63024826579043	3.84585323491099	2.59516130035904	2.67018506176497	2.19678873653315	
14000US01003011602	4	2.77117296222664	10.4201400466822	7.69230769230769	2.72818428184282	3.23687752355316	2.13762052753207	2.26639220078572	2.09398928204696	
14000US01007010002	3	2.24420632942935	21.6546762589928	35.5371900826446	2.58518673940411	3.16041848299913	2.19503787686207	2.41208734706003	2.05519668989453	
14000US01007010003	2	2.60460652591171	20.685332218972	0	2.68467027559055	4.64189189189189	2.19633353045535	2.49248591876806	2.21600990733529	
14000US01007010004	2	2.16322104958797	15.6236559139785	0	2.13385199705954	3.03169213355971	2.48540942526877	2.07318680700586	1.89408364307356	
14000US01009050101	2	2.43017843289372	22.3537414965986	32.2033898305085	2.48763866877971	4.1198808637379	2.1807807711951	2.56679399828179	2.16216404637883	
14000US01009050102	2	2.41663342640606	19.674039580908	37.5	2.63722730471499	3.73282442748092	2.47613697164259	2.59163061891348	2.10179750482119	
14000US01009050300	2	2.25860597439545	21.4125788723794	0	2.33577430972389	2.55873925501433	2.16462608873411	2.02904146379383	2.05755456464531	

# **Appendix - Dashboard View by US Ignite**

Dropdown menu to select tracts and apply filters usignite Data tools Broadband 49047968200 Uintah, UT Visualizations Correlation socioeconomic and broadband score Filter Broadband score Socioeconomic Factors State County Filter presets Show/hide layers Download data ↓ Education level over the years Education level Socioeconomic factor Broadband score 2016 2017 2018 2019 2022 2023 1 2 3 4 5

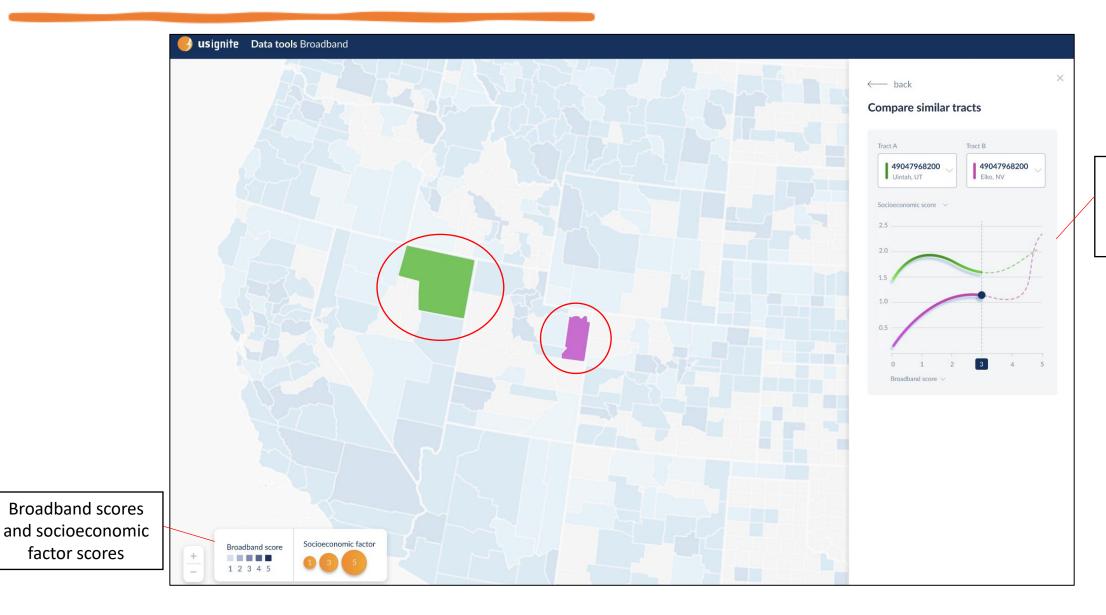
Expected growth from increased broadband score

Changes in scores over time

Broadband scores and socioeconomic factor scores

# **Appendix - Comparing Similar Tracts**

factor scores



Comparison of expected growth from increased broadband score

37

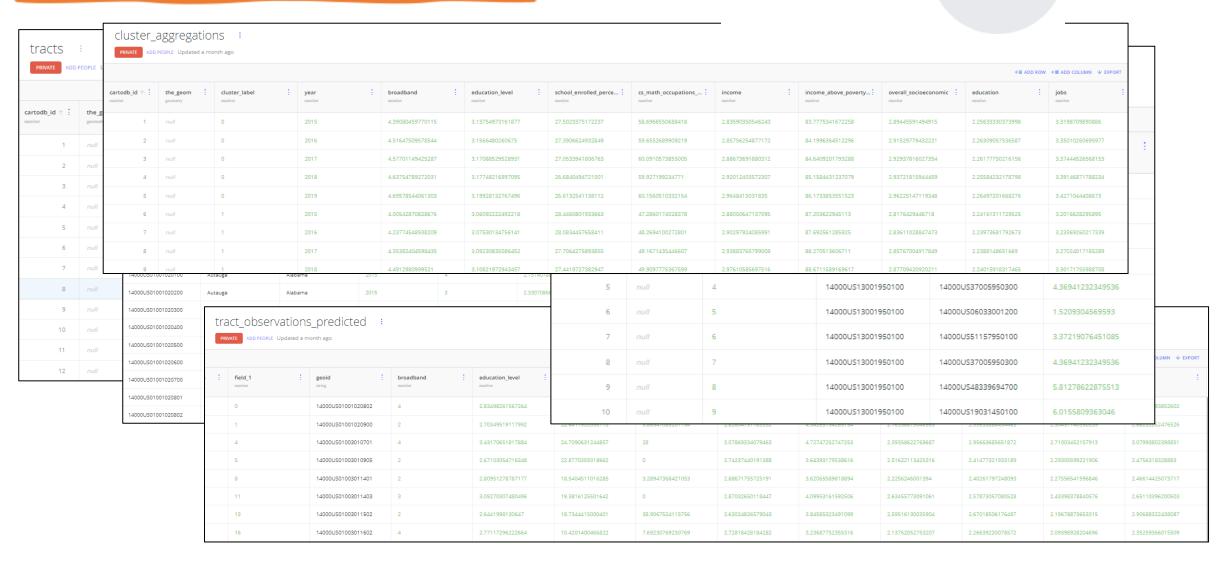
# **CARTO – Data Tables**



TABLE	DESCRIPTION
tracts	Census tract level information
tract_observations/county/state	Census tract level information with socioeconomic scores
tract_observations_predicted/county/state	Predicted socioeconomic scores
tracts_with_closest_neighbours	Top 5 similar tracts
cluster_aggregations	Output of 12 distinct clusters based on the features of the areas. Values are average of all years observed

# **CARTO – Data Tables**

# CARTO



# PROJECT COMPLETION PLAN

Project Activities / Week#	1	2	3	4	5	6	7	8	9	10	Summer Break	1	2	3	4	5	6	7	8	9	10
Project and Research Design																					
Client / Advisor Introductions																					
Finalize Project Scope																					
Project Implementation										$\Rightarrow$											
New Data Collection / EDA																					
Understand Clusters and Modeling																					
Validate Previous Models with New Data											2 – 3 checkpoint meetings / emails										
Segment Communities Across Time Periods											with US Ignite and project advisor during Summer										
Identify Factors that Influence Change (Telling the Story)											Break										
Research Paper Draft										$\Rightarrow$											
Create Case Templates by Community Segments																					
Validate Cases																					
Project Delivery																					
Capstone Showcase																					





# THANK YOU!

