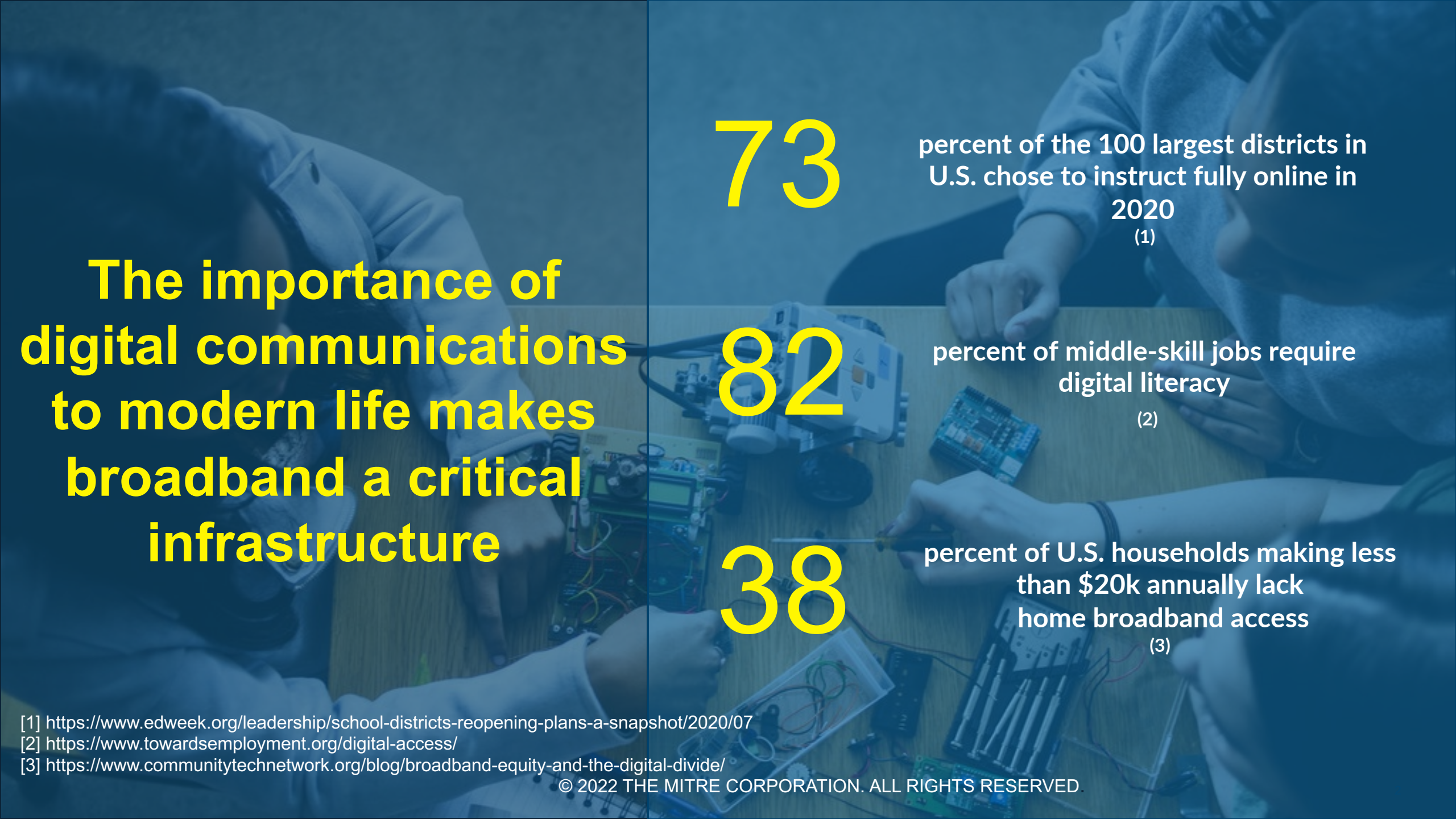
Market Segmentation of State Broadband Delivery Approaches

’22 ACM/IEEE Next G Summit Technology Track

Anna Grabek
Silvia Ionescu
Carolyn Kahn

MITRE | SOLVING PROBLEMS FOR A SAFER WORLD™



The importance of digital communications to modern life makes broadband a critical infrastructure

73 percent of the 100 largest districts in U.S. chose to instruct fully online in 2020 (1)

82 percent of middle-skill jobs require digital literacy (2)

38 percent of U.S. households making less than $20k annually lack home broadband access (3)

[1] https://www.edweek.org/leadership/school-districts-reopening-plans-a-snapshot/2020/07
[2] https://www.towardsemployment.org/digital-access/
[3] https://www.communitytechnetwork.org/blog/broadband-equity-and-the-digital-divide/

The Infrastructure Investment and Jobs Act of 2021 is making more than $100 million available to each state for broadband improvement

- Investment and allocative decisions to be made at state level
- Understanding current state broadband delivery models informs potential challenges and opportunities
- Decisions made now will have long-term impacts on Next G technologies

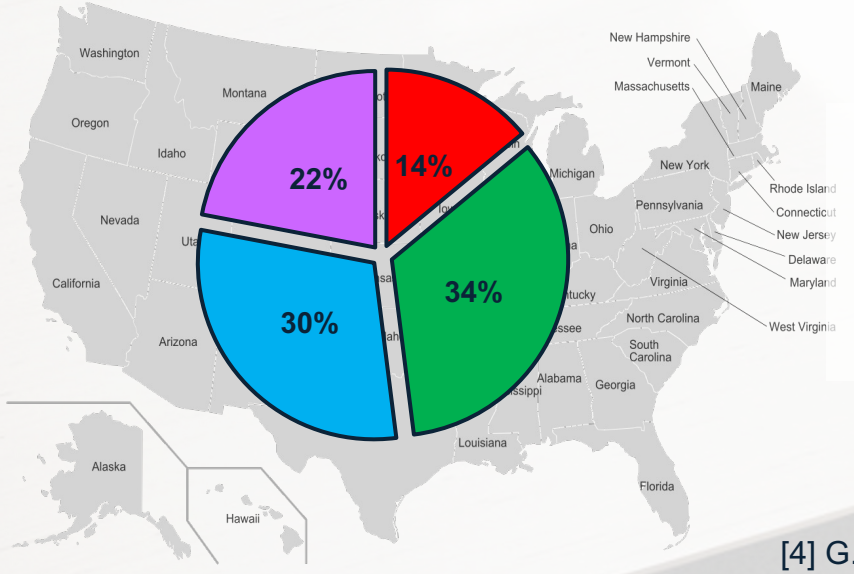
Understanding state broadband delivery approaches is desirable, but inherently challenging due to:

- Non-uniformity of approaches
- Trade-offs between complexity and utility
- Unobservable factors

Methodology

Background

A *market segmentation* is typically performed for the purpose of understanding the common profiles that exist in a target market of interest [4]



Segmenting the U.S. states allows one to understand the broadband delivery profiles that exist

[4] G. Martin, “The importance of marketing segmentation” American Journal of Business Education, vol. 4, no. 6, 2021.

Dataset

Dataset describes three general dimensions of broadband delivery, using publicly available data on state broadband programs

Administrative Resources Data:

- Presence of central broadband office
- Number of broadband agencies
- Number of broadband task forces

Strategic Planning Data:

- Whether the state has developed an official broadband strategy and if so, in what year
- Whether the state has developed an official broadband goal and if so, in what year

Execution Data:

- Whether the state has current mapping of areas served/ underserved
- Number of state broadband funds
- Number of state laws pertaining to broadband

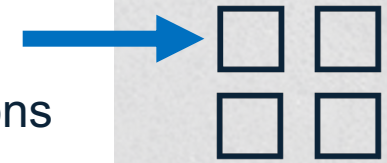
Segmentation Process

Segmentation input is data on state broadband programs

Segmentation Process

Record Data	Quantify presence and extent of broadband programs and resources for each state. If appropriate, convert cardinal data points to binary data indicating category or relationship to median value of data point in set • Ex: cardinal 'number of laws' becomes binary feature 'has less than 5 broadband laws' or 'has greater than median laws'
Dimensionality Reduction	Transform dataset into lower dimensional space so that clustering is more tractable
Cluster	Group states according to their commonalities along given dimensions

Segmentation output is states categorized by broadband delivery profile



Dimensionality Reduction

Feature dataset reduced to 3-dimensional space using two algorithms:

- Apriori Algorithm: Derive association rules amongst features using measures of conditional dependence, then score states based on associations with leading indicators

  P(Consequent Set | Antecedent Set)

- Linear Regression: Predict leading indicator scores for states using remaining features in set

$$\hat{y} = b_0 + \sum_{i=1}^{N} b_i x_i$$

Reduced datapoints correspond to "scores" along foundational dimensions of broadband delivery

K-Means Clustering

Dataset of state broadband program scores

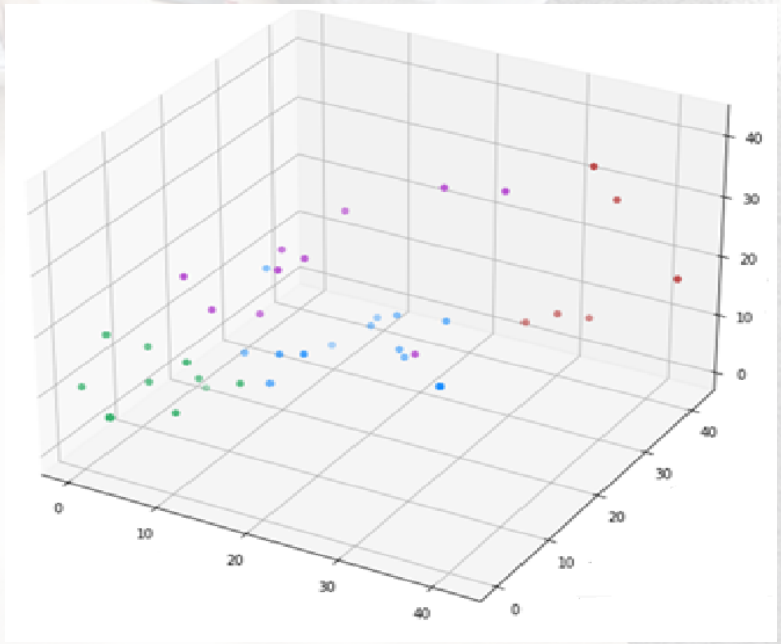
Determine Number of Clusters

Initialize Cluster Centers

Decide Class Membership

Re-estimate Cluster Centers

Unstable clusters



Mixed Clustering with K-Medoids and Gower's Distance

Dataset of raw state broadband program features

Derive Gower Distance Matrix

Determine Number of Clusters

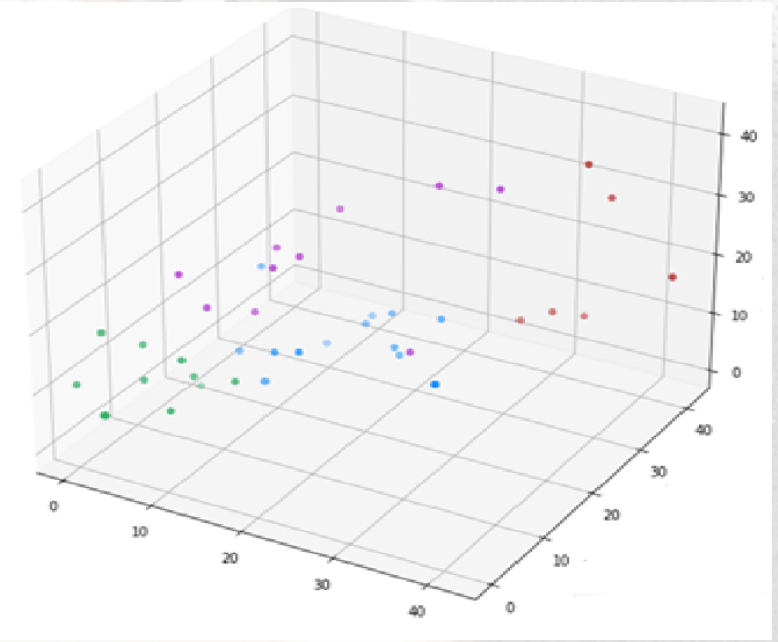
Derive Cluster Medoids

Decide Class Membership

Re-estimate Cluster Medoids

Unstable clusters

Gower's distance is a metric used to measure the similarity between two data points that contain both numeric and categorical variables [5]

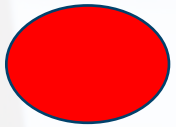


[5] Gower, J. C. (1971). A general coefficient of similarity and some of Its properties.

Results

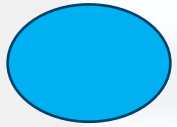
General Trends

Four general broadband delivery profiles observed across segmentations

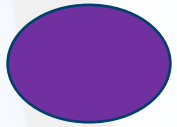


States with low evidence of activity or capacity along all dimensions of broadband delivery program

States with high evidence of activity and capacity along all dimensions of broadband delivery program



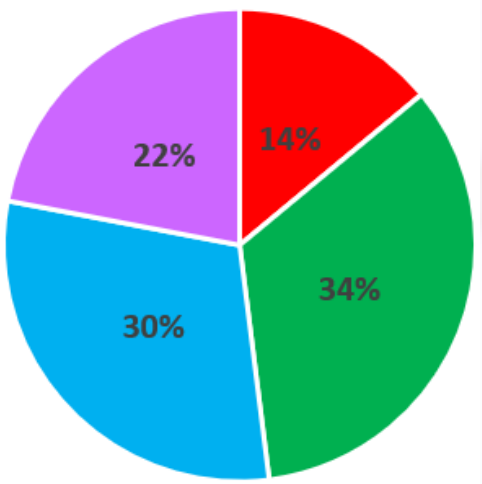
States with predominantly low evidence of administrative organization



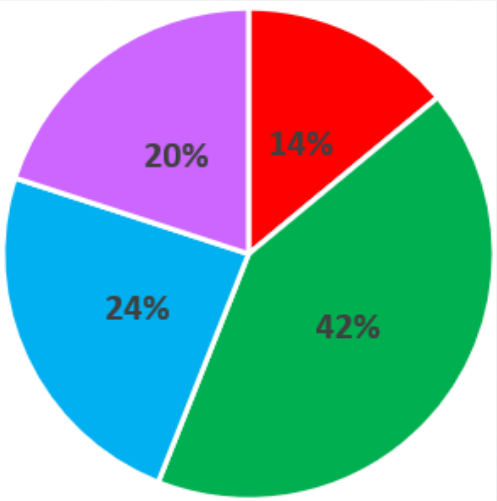
States with predominantly low evidence of strategic activity

Segmentations

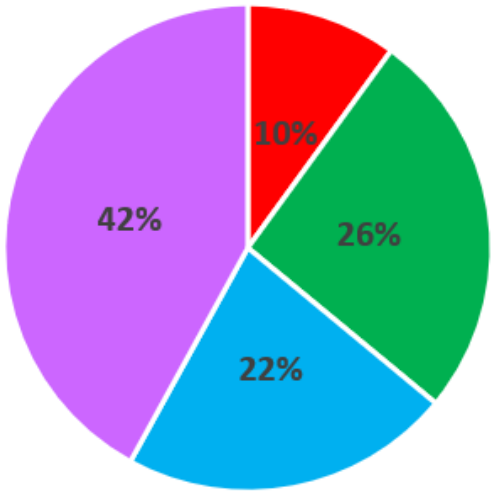
Apriori with K-Means and categorical binary feature set



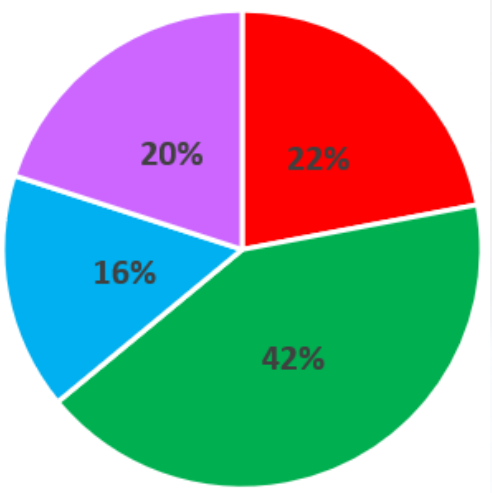
Apriori with K-Means and median binary feature set



Linear Regression with K-Means Clustering



Mixed Clustering



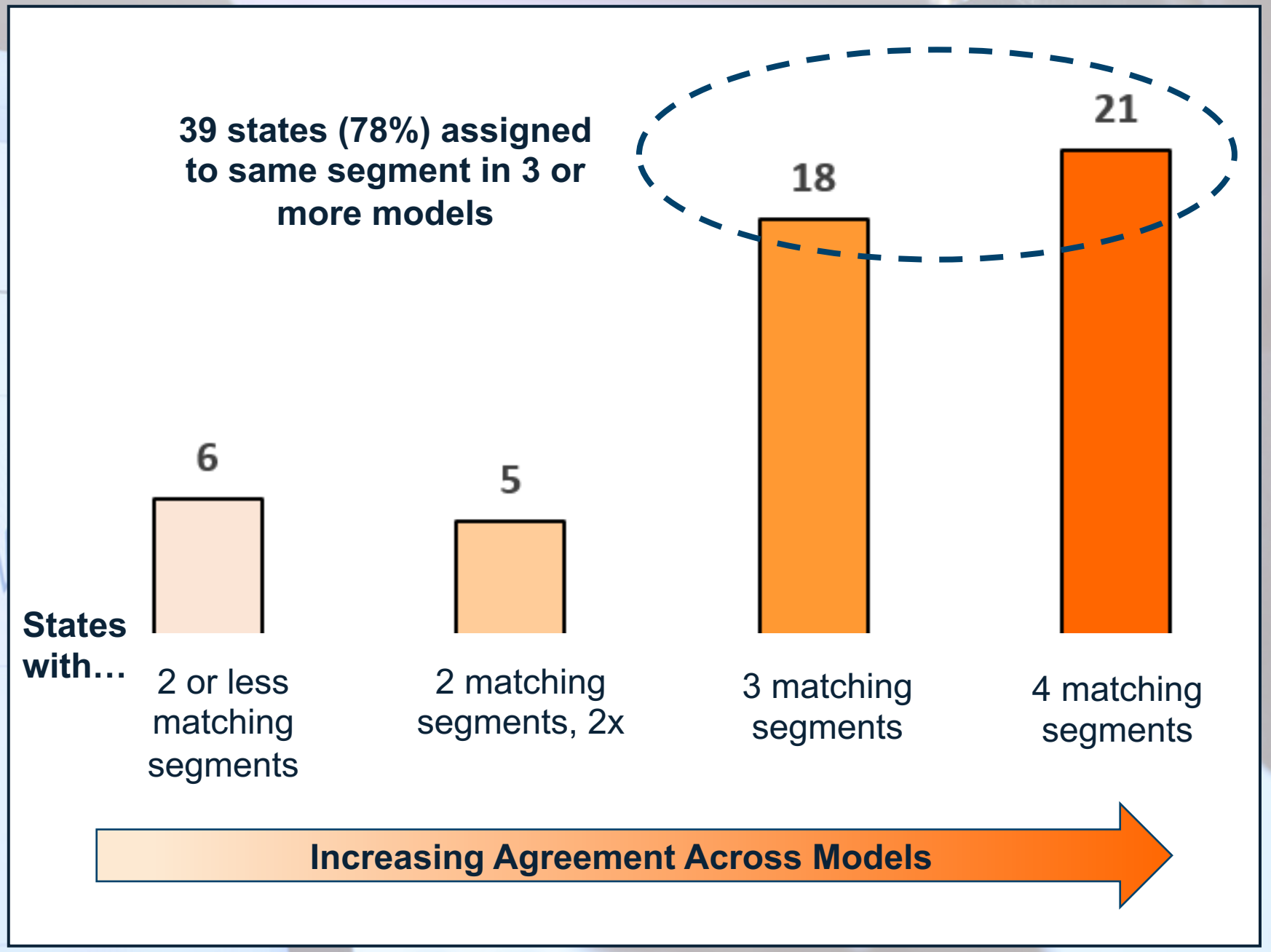
Low evidence in all dimensions

High evidence in all dimensions

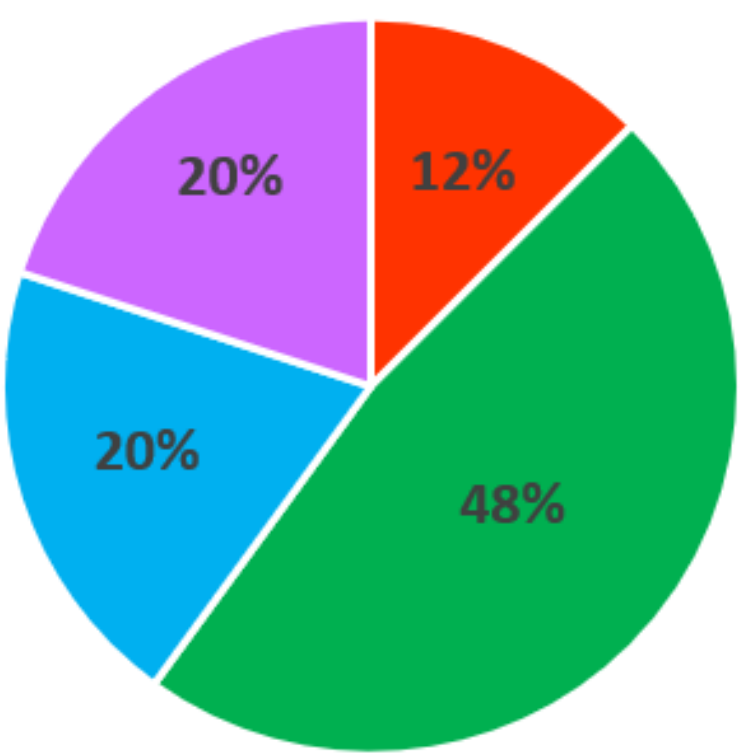
Low evidence of administrative org.

Low evidence of strategic activity

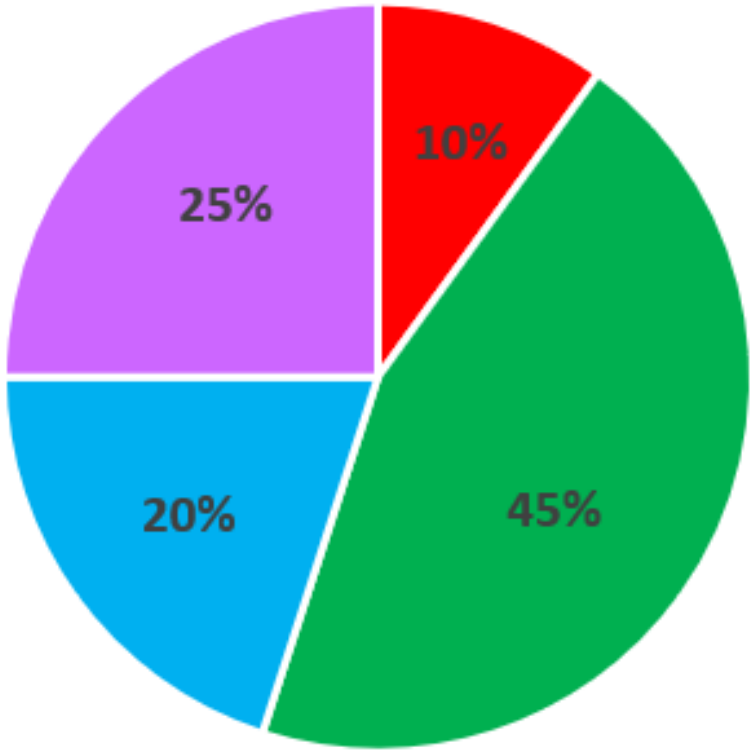
Consensus



Consensus



39 states assigned to same segment in 3 or more models



21 states assigned to same segment in all 4 models

Low evidence in all dimensions

High evidence in all dimensions

Low evidence of administrative org.

Low evidence of strategic activity

Discussion

The Impact on Next G

Broadband improvements and investments made by states now will shape the U.S. Next G landscape

Anna Grabek agrabek@mitre.org
Silvia Ionescu sionescu@mitre.org
Carolyn Kahn ckahn@mitre.org