



# Broadband mapping meeting with FCC

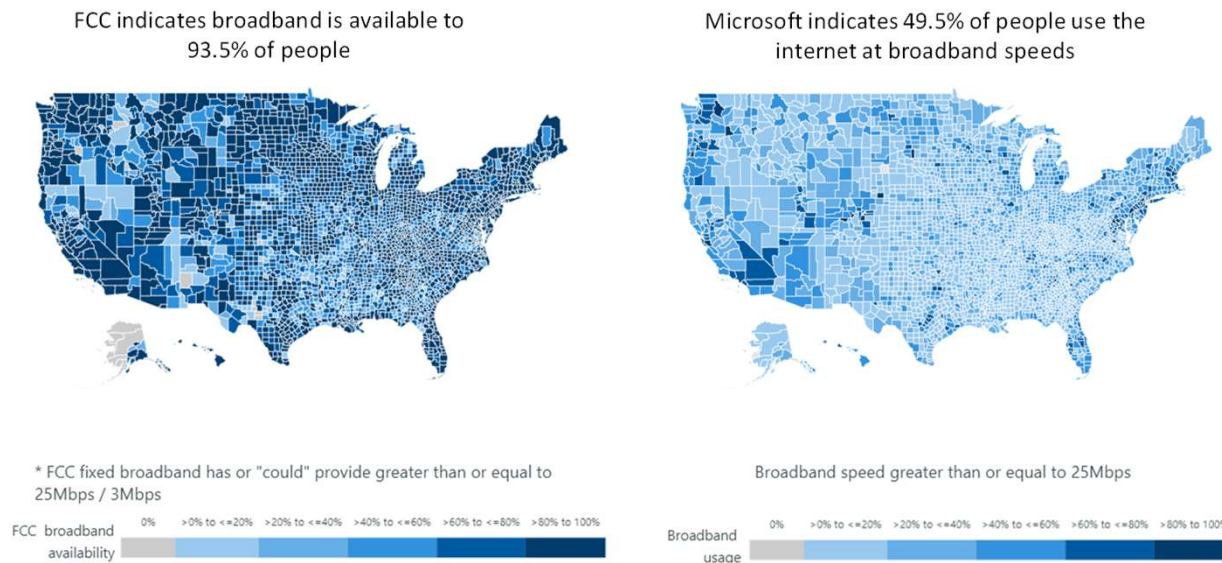
September 18, 2019

## Areas of discussion

- Form 477 vs Digital Opportunity Data Collection order
- Outlier zip code analysis on existing availability data

# Broadband usage based on Microsoft data

- FCC reports 93.5% of the country has access to fixed broadband at a minimum of 25 Mbps/3Mbps; Microsoft estimates ~49% of people access the internet at broadband speeds
  - Availability does not equal usage; however usage gives us the ground truth in the progress we are making in broadband adoption.
  - Through artificial intelligence and machine learning models using device level (no PII) data (over 200+ Microsoft services) we estimate download speeds and broadband coverage
  - We make a very minor adjustment in areas of the country that Microsoft may not have a presence with third party data i.e. ComScore



Data sources: FCC 2019 Broadband report based on Form 477 data from December 2017 and Microsoft data from September 2018

## Objective of the analysis

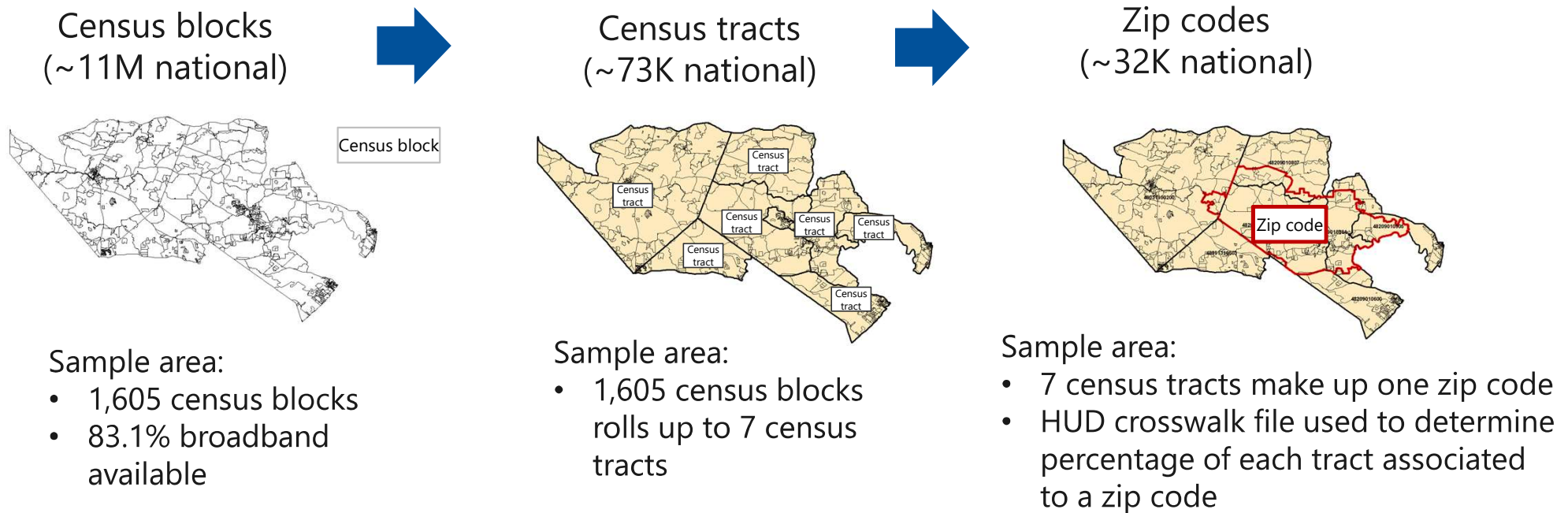
- Hypothesis: If we can find zip codes with inaccuracies in availability data in an automated way using machine learning this could help the stakeholders to correct data inaccuracies.
- Utilizing further machine learning to predict availability, we have created a model to identify a subset of zip codes that MAY have inaccuracies.
- There is no guarantee that these zip codes are being reported inaccurately; however based on using a machine learning model and additional validation with a third-party survey done by BroadbandNow, these identify areas of potential inaccuracies.
- Our plan is to make the model publicly available on GitHub and the output publicly available in the near future.

# Methodology

- Developed Machine Learning models (random forest for regression) to predict broadband availability in order to identify potential outlier zip codes when compared to the form 477 data submitted to the FCC.
- We take the FCC availability data at the census tract level and estimate to zip codes.
- Data sources:
  - FCC Form 477 (grouped by zip code)
  - Broadband usage based on Microsoft data
  - Census data by zip code
  - Broadbandnow.com data
  - HUD census tract to zip code crosswalk

# Zip code availability

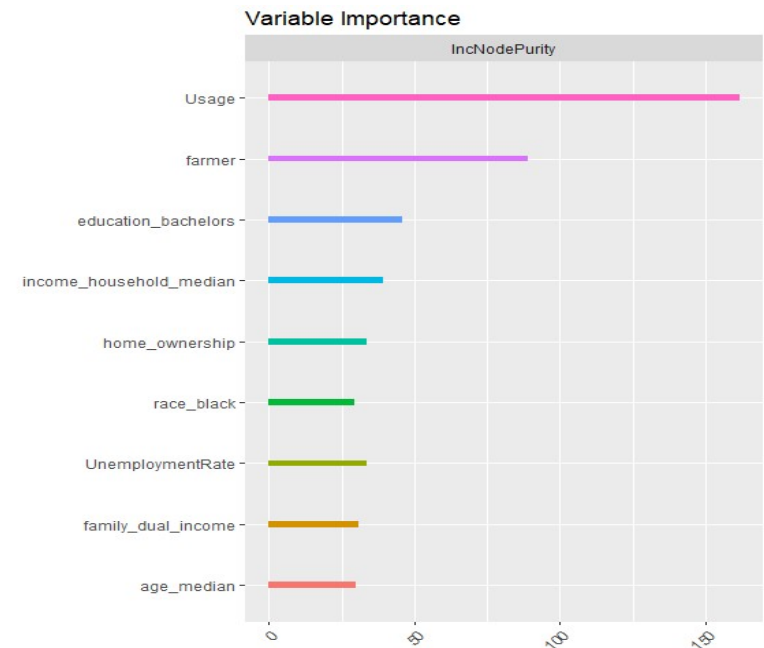
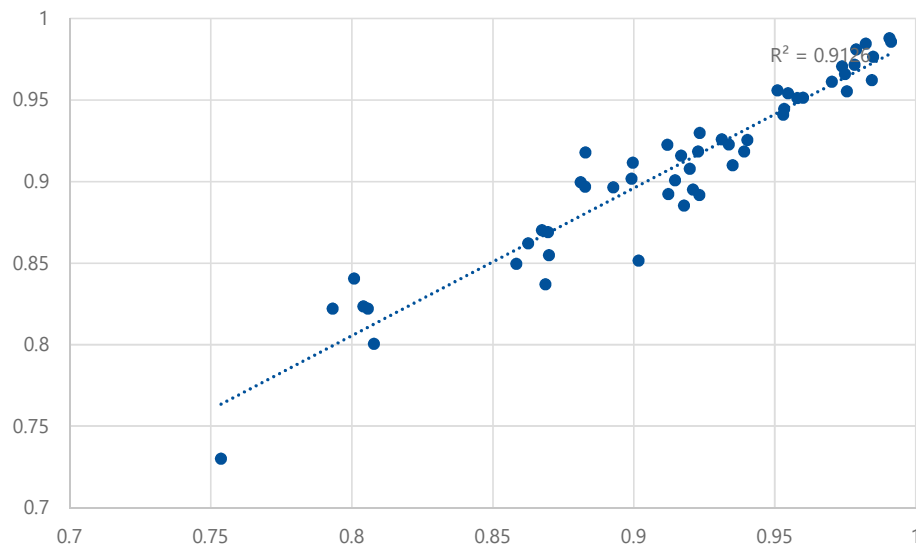
- We take the FCC availability data at the census tract level and estimate to zip codes using the HUD crosswalk file



# Predicting broadband availability

- We use a random forest model and measure variable importance.
- Broadband usage is the #1 variable with the highest predictive power followed by percent of farmer and educational attainment.
- At the state level this model can predict with an  $r^2$  of 91%.

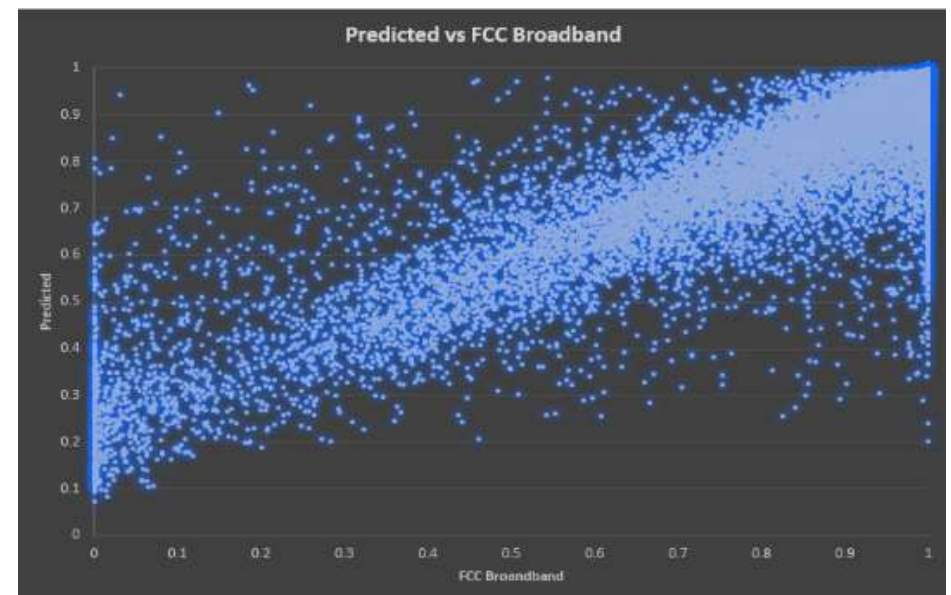
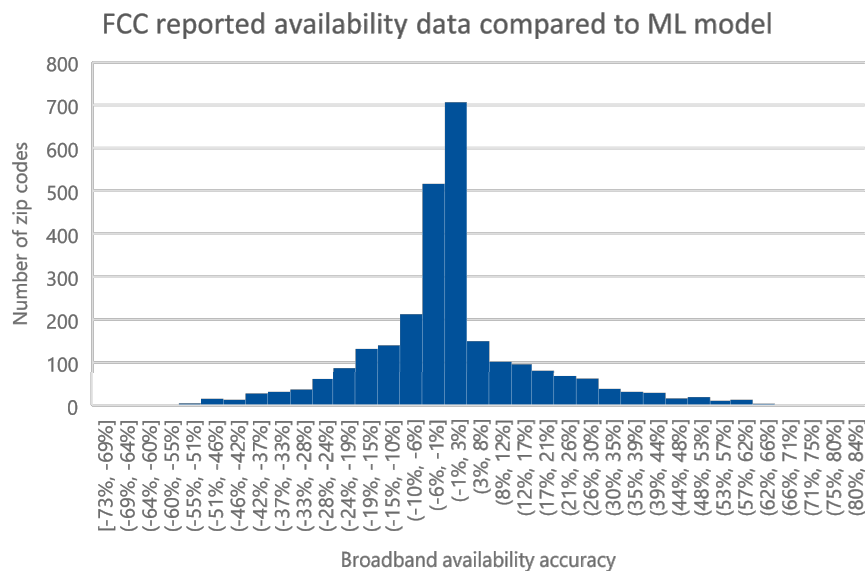
Predicted Broadband vs FCC broadband (State Level)





# Ability to detect potential outliers

- We use this model to detect potential outliers with the highest divergence (positive and negative) to the reported broadband availability
  - ML model predicts 63% of zip codes within 5pts of reported broadband availability
  - ML model predicts 78% of zip codes within 10pts of reported broadband availability





## Top 20 potential outlier zip codes

State	Zip code	FCC broadband availability 2019 report	Usage Feb 2019
PA	17949	91.8%	0.0%
VA	22742	100.0%	0.5%
WV	26386	100.0%	9.5%
FL	33890	94.0%	4.7%
OH	44076	92.1%	5.9%
OH	45856	98.3%	4.5%
IA	50514	98.0%	4.5%
MN	56282	100.0%	3.7%
KS	66079	100.0%	0.6%
AR	71956	97.3%	6.1%
AR	71968	99.1%	7.7%
OK	74332	99.9%	0.7%
TX	78118	100.0%	3.4%
TX	78151	99.6%	0.5%
TX	78941	99.5%	2.8%
CA	93602	93.7%	8.5%
CA	95638	100.0%	2.2%
OR	97456	94.7%	7.6%
WA	98855	97.9%	7.4%
WA	99122	100.0%	4.0%

# Zip code: 17949 in Pennsylvania



Estimated FCC broadband availability\*  
(2019 report)

91.8%

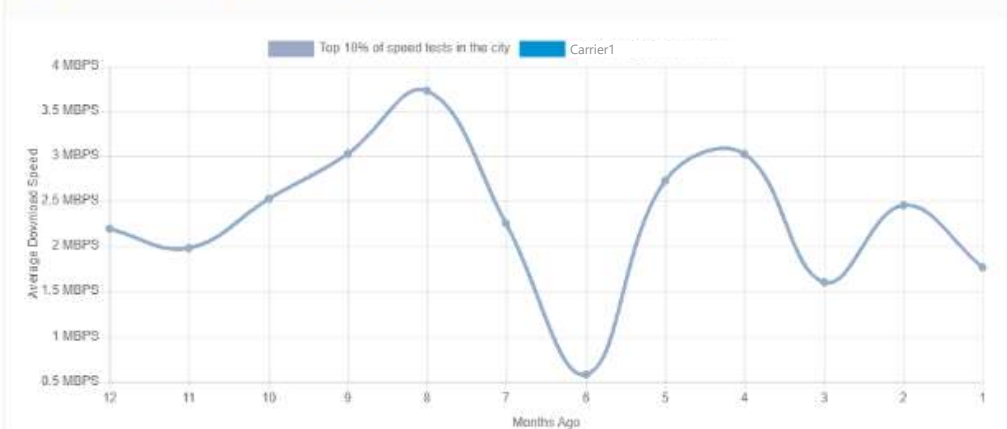
Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

0.0%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

DOWNLOAD SPEEDS IN MAHANOY PLANE



This analysis is based on 336 speed tests from IP verified users who took speed tests from an IP address in Mahanoy Plane between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.

Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com

# Zip code: 22742 in Virginia



Estimated FCC broadband availability\*  
(2019 report)

100.0%

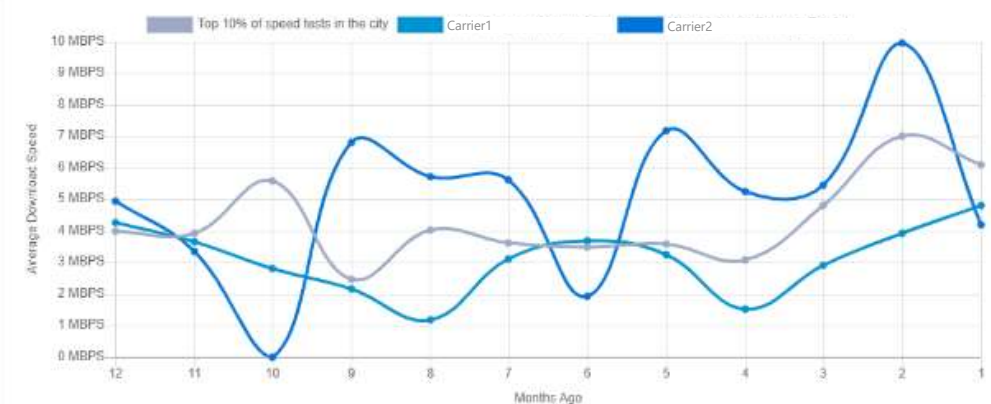
Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

0.5%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

DOWNLOAD SPEEDS IN SUMNERDUCK

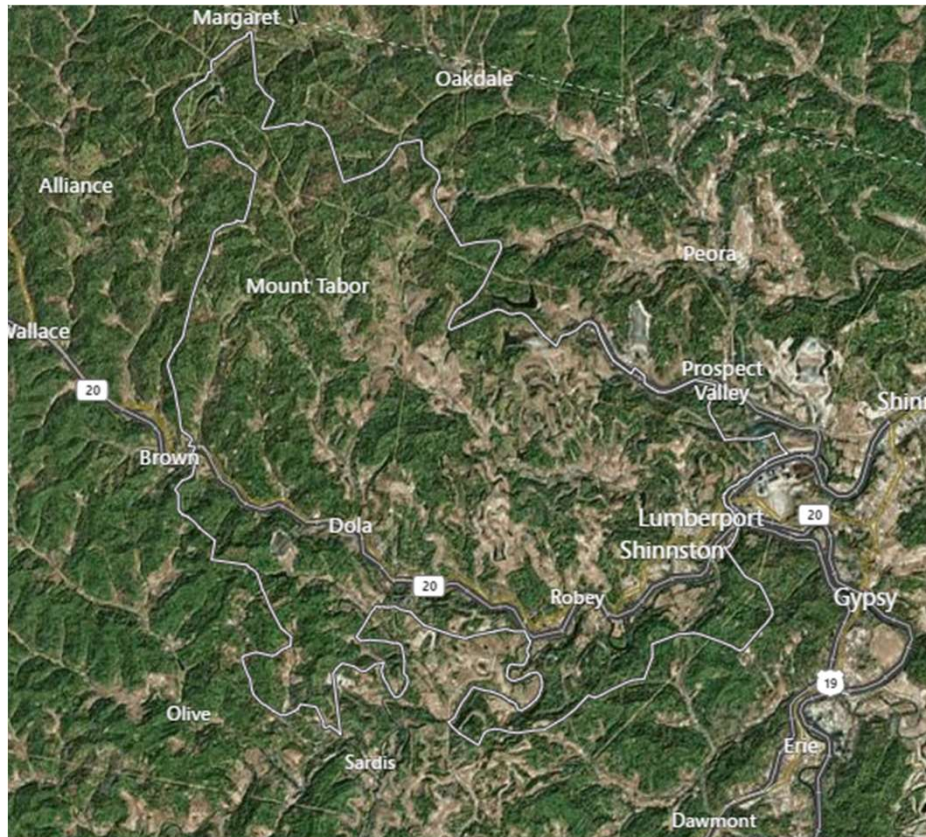


This analysis is based on 638 speed tests from IP-verified users who took speed tests from an IP address in Sumnerduck between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.

Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com



# Zip code: 26386 in West Virginia

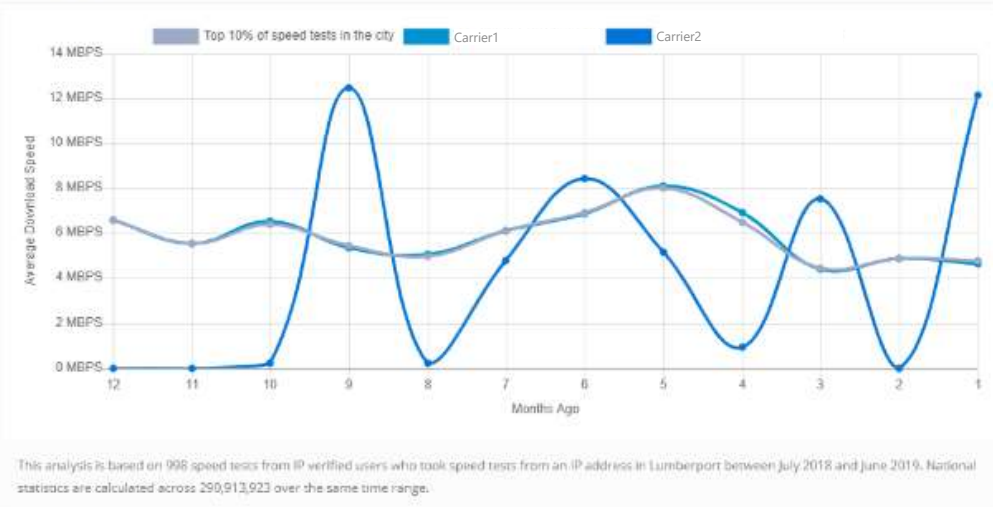


Estimated FCC broadband availability* (2019 report)	Estimated percent of people using the internet at broadband speeds using Microsoft data
100.0%	9.5%

\* zip codes may contain portions of multiple census tracts

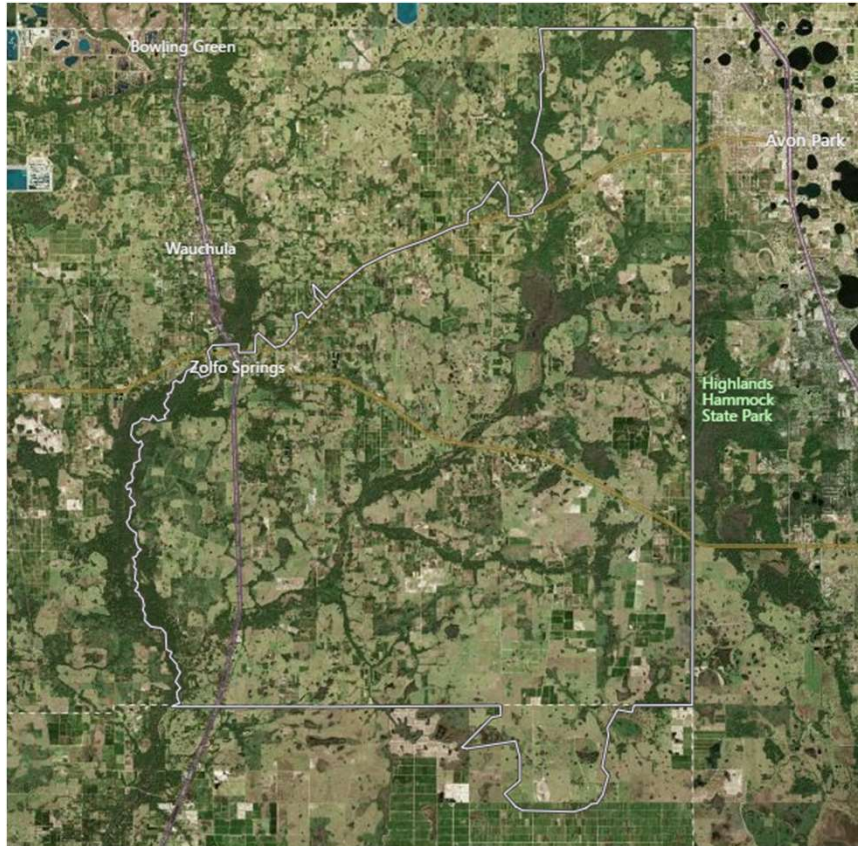
BROADBANDNOW®

DOWNLOAD SPEEDS IN LUMBERPORT



Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com

# Zip code: 33890 in Florida



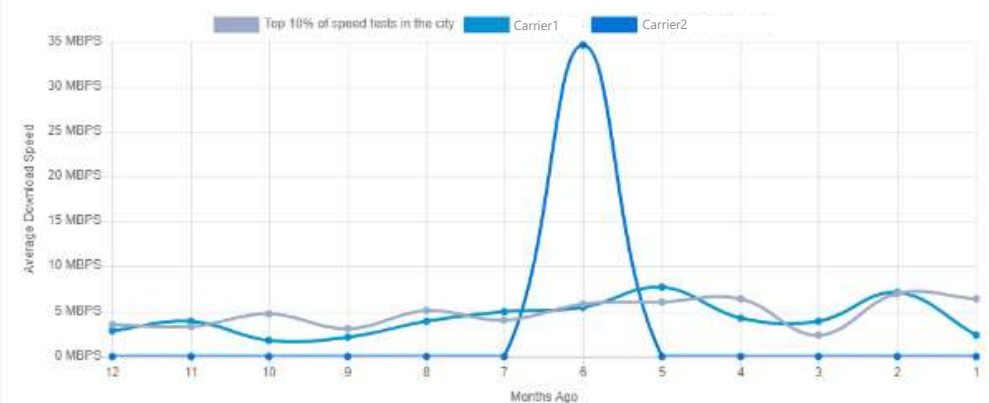
Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com

Estimated FCC broadband availability* (2019 report)	Estimated percent of people using the internet at broadband speeds using Microsoft data
94.0%	4.7%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

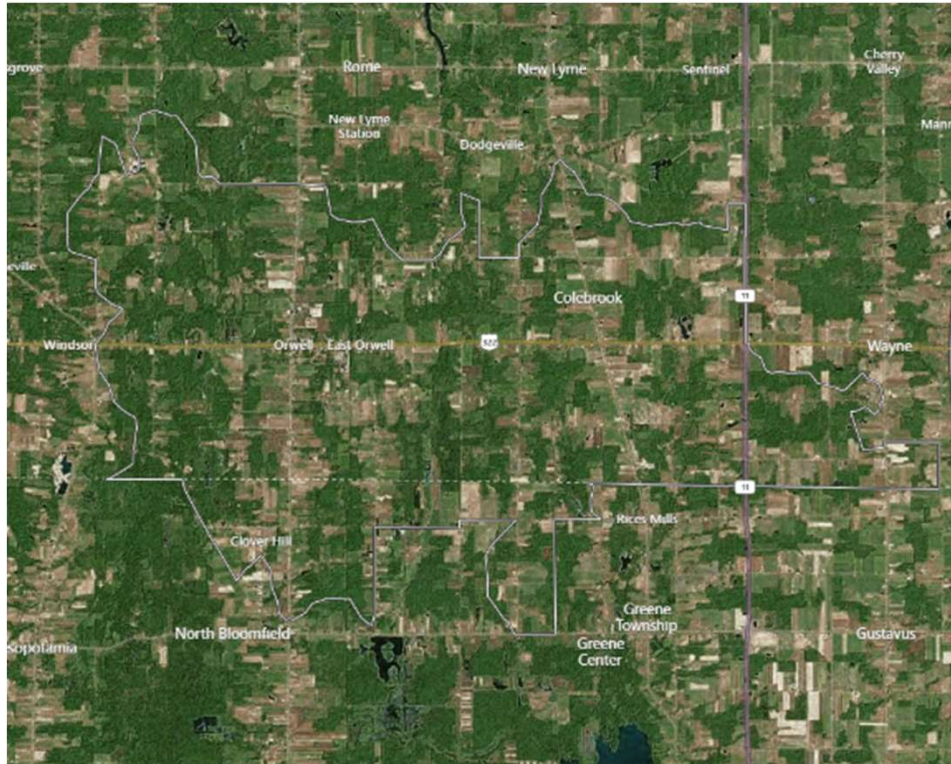
## DOWNLOAD SPEEDS IN ZOLFO SPRINGS



This analysis is based on 856 speed tests from IP verified users who took speed tests from an IP address in Zolfo Springs between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.



# Zip code: 44076 in Ohio



Estimated FCC broadband availability\*  
(2019 report)

92.1%

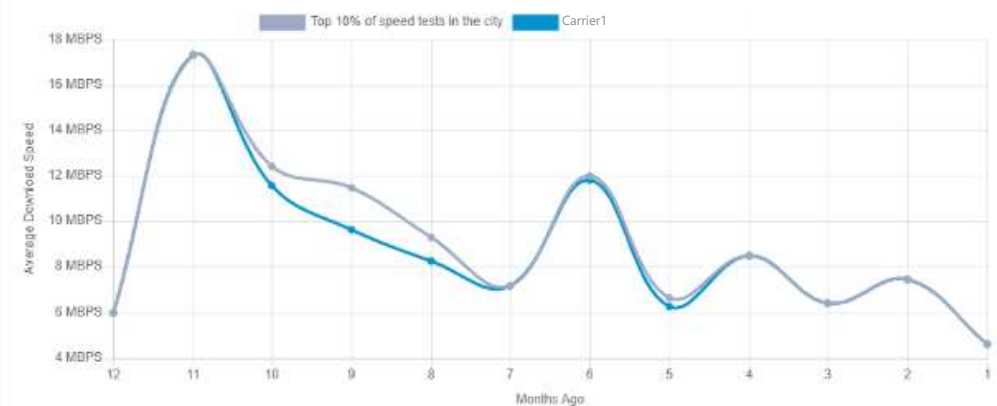
Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

5.9%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

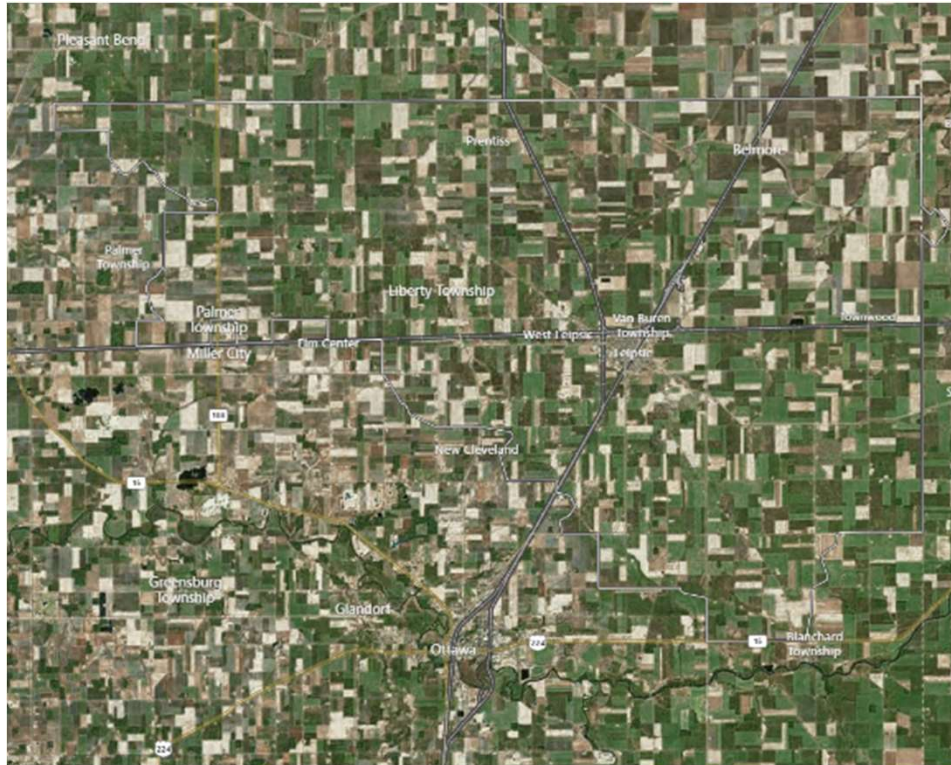
DOWNLOAD SPEEDS IN ORWELL



This analysis is based on 1,690 speed tests from IP verified users who took speed tests from an IP address in Orwell between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.

Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com

# Zip code: 45856 in Ohio



Estimated FCC broadband availability\*  
(2019 report)

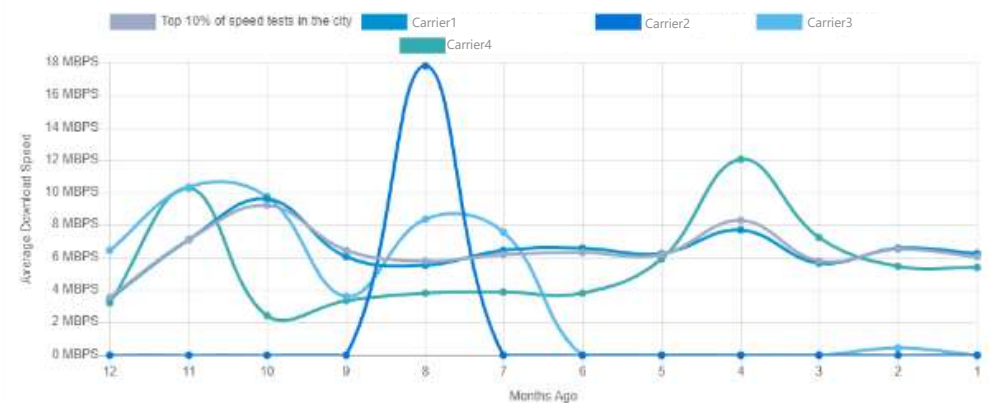
98.3%

Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

4.5%

\* zip codes may contain portions of multiple census tracts

## BROADBANDNOW® DOWNLOAD SPEEDS IN LEIPSIC

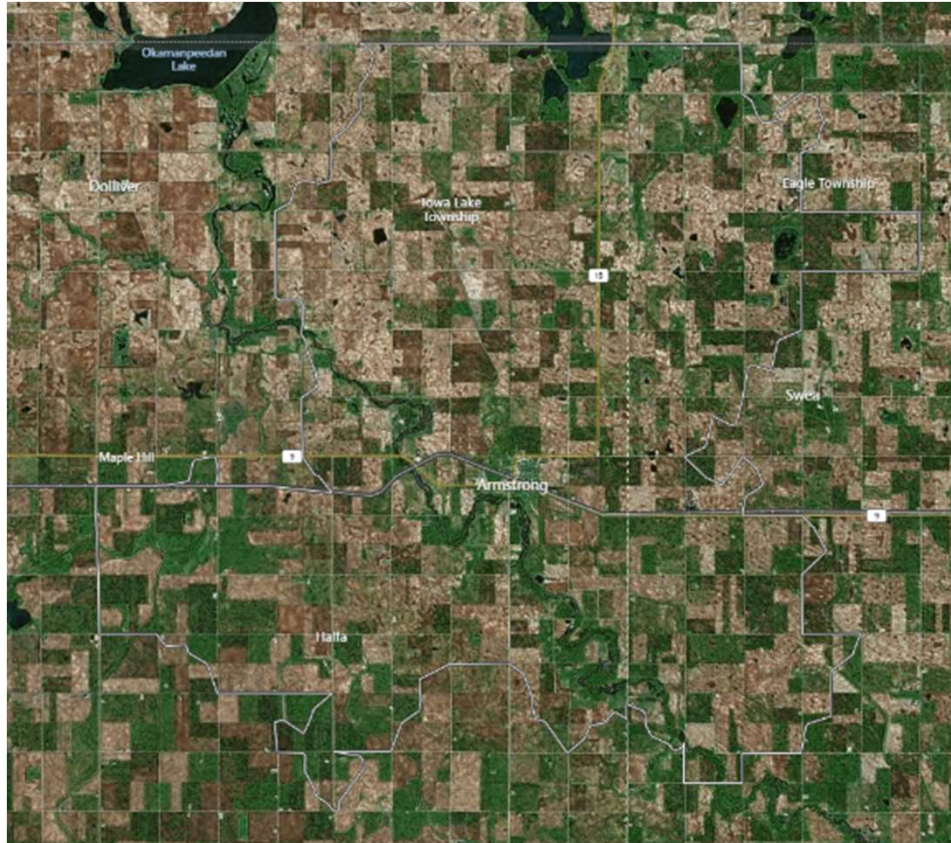


This analysis is based on 3,003 speed tests from IP verified users who took speed tests from an IP address in Leipsic between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.

Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com



# Zip code: 50514 in Iowa



Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com

Estimated FCC broadband availability\*  
(2019 report)

98.0%

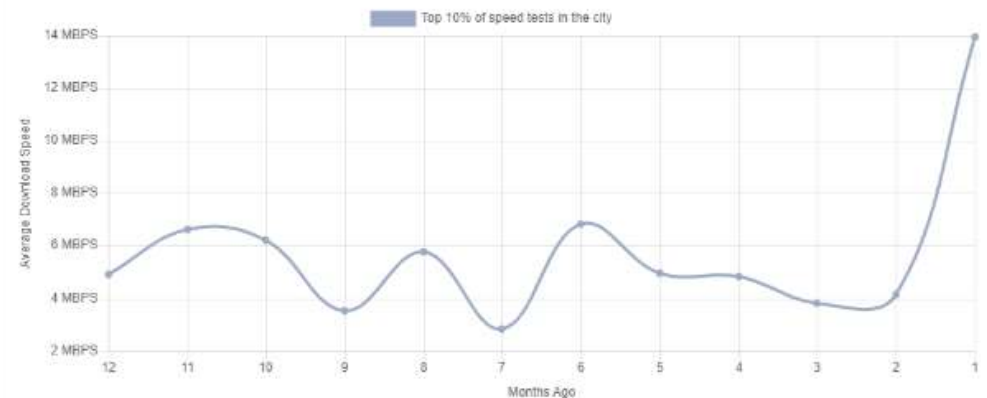
Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

4.5%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

DOWNLOAD SPEEDS IN ARMSTRONG



This analysis is based on 1,084 speed tests from IP verified users who took speed tests from an IP address in Armstrong between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.

# Zip code: 56282 in Minnesota



Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com

Estimated FCC broadband availability\*  
(2019 report)

100.0%

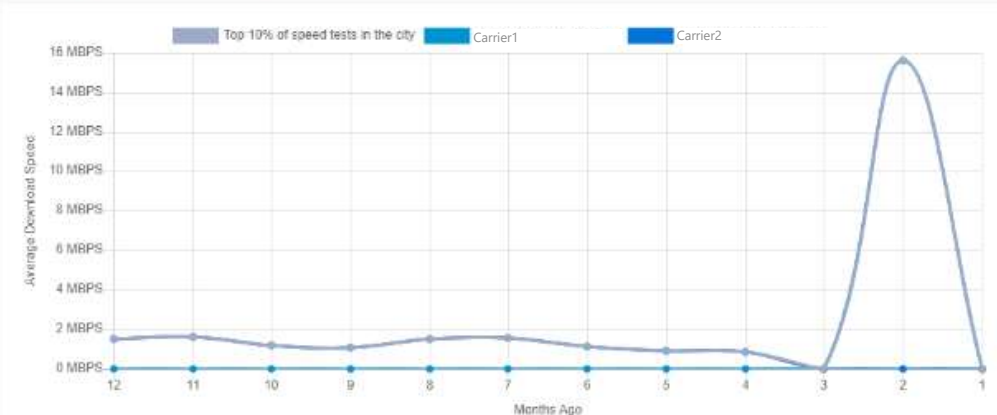
Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

3.7%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

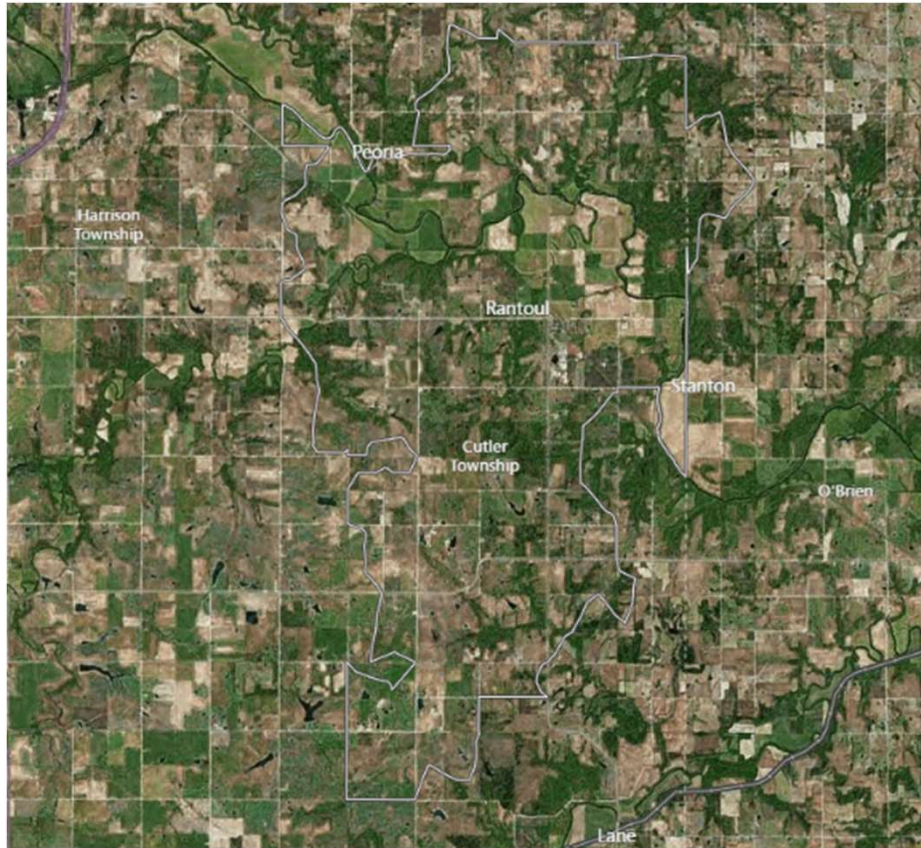
DOWNLOAD SPEEDS IN RAYMOND



This analysis is based on 344 speed tests from IP verified users who took speed tests from an IP address in Raymond between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.



# Zip code: 66079 in Kansas



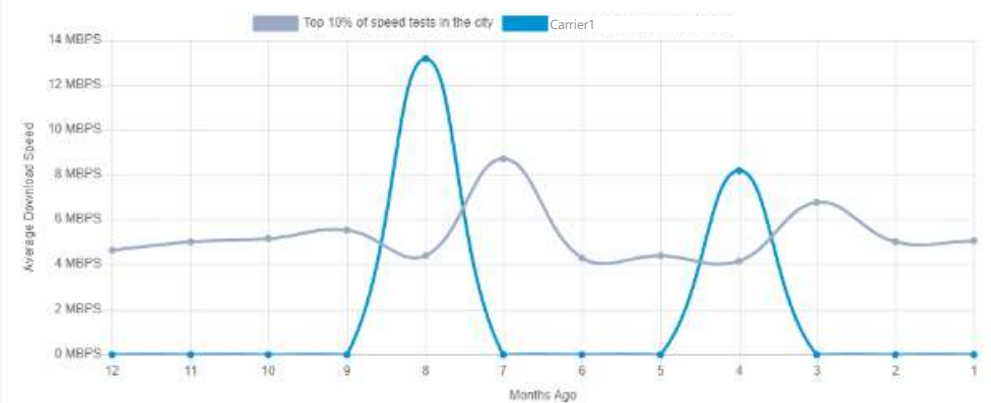
Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com

Estimated FCC broadband availability* (2019 report)	Estimated percent of people using the internet at broadband speeds using Microsoft data
100.0%	0.6%

\* zip codes may contain portions of multiple census tracts

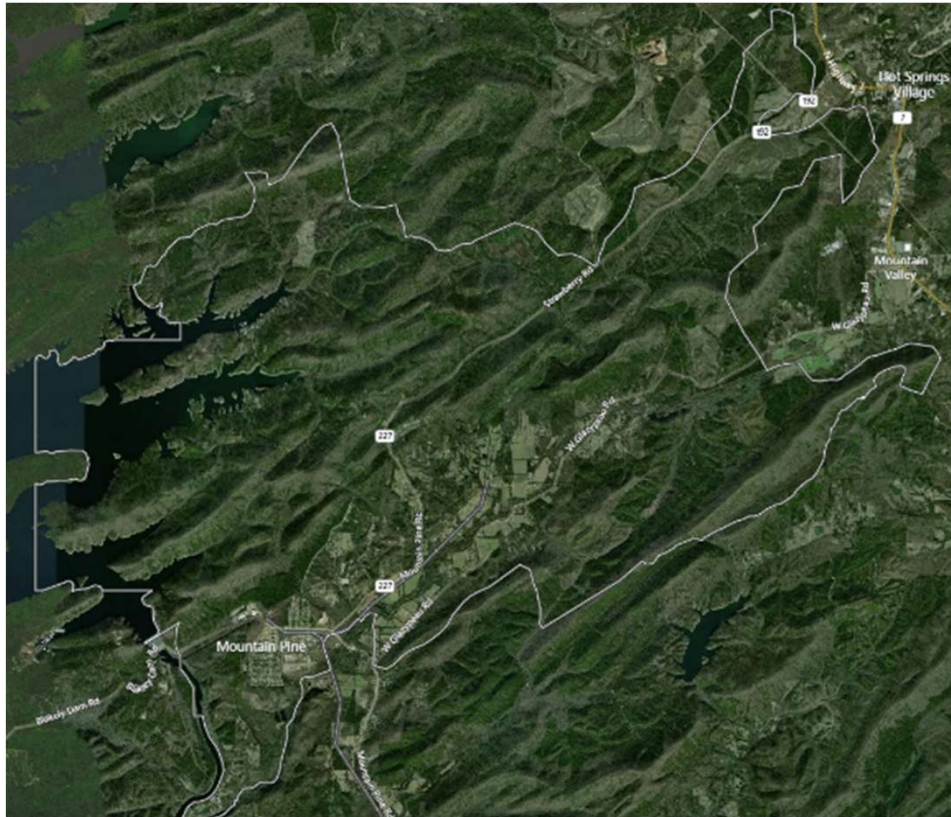
## BROADBANDNOW®

### DOWNLOAD SPEEDS IN RANTOUL



This analysis is based on 1,293 speed tests from IP verified users who took speed tests from an IP address in Rantoul between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.

# Zip code: 71956 in Arkansas



Estimated FCC broadband availability\*  
(2019 report)

97.3%

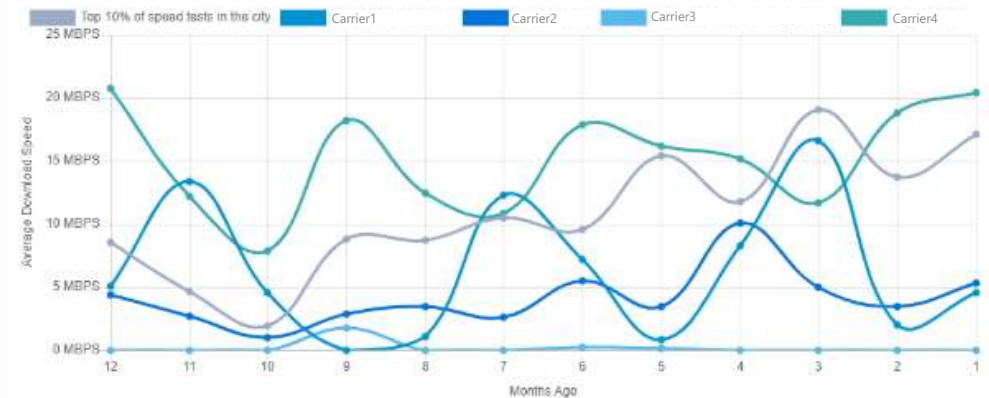
Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

6.1%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

DOWNLOAD SPEEDS IN MOUNTAIN PINE

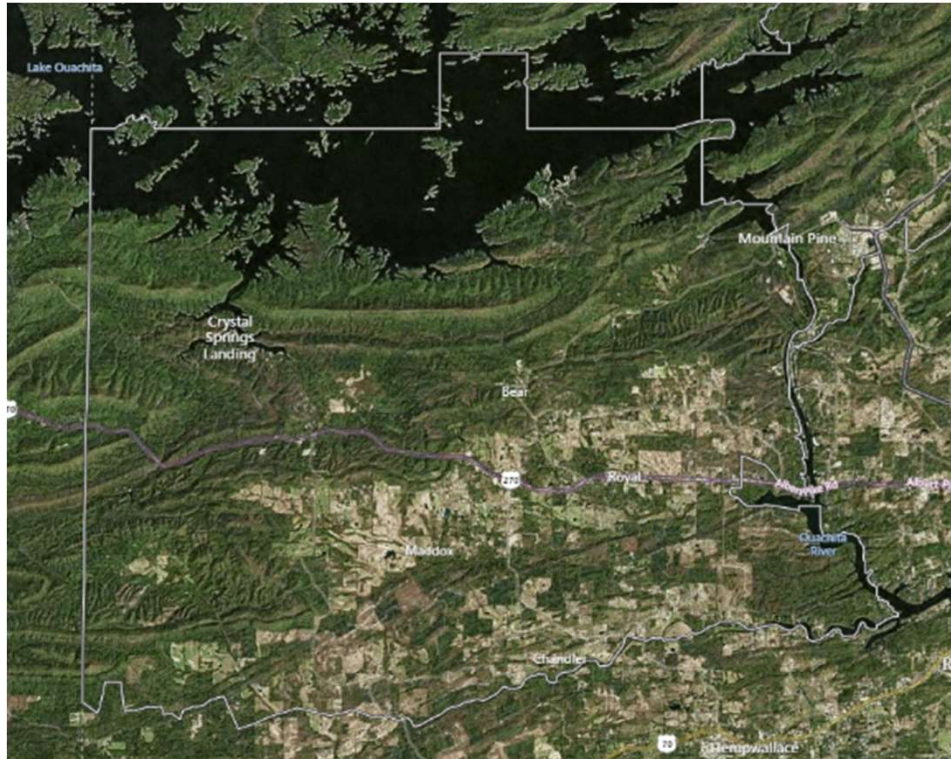


This analysis is based on 1,512 speed tests from IP verified users who took speed tests from an IP address in Mountain Pine between July 2018 and June 2019. National statistics are calculated across 250,913,923 over the same time range.

Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com



# Zip code: 71968 in Arkansas



Estimated FCC broadband availability\*  
(2019 report)

99.1%

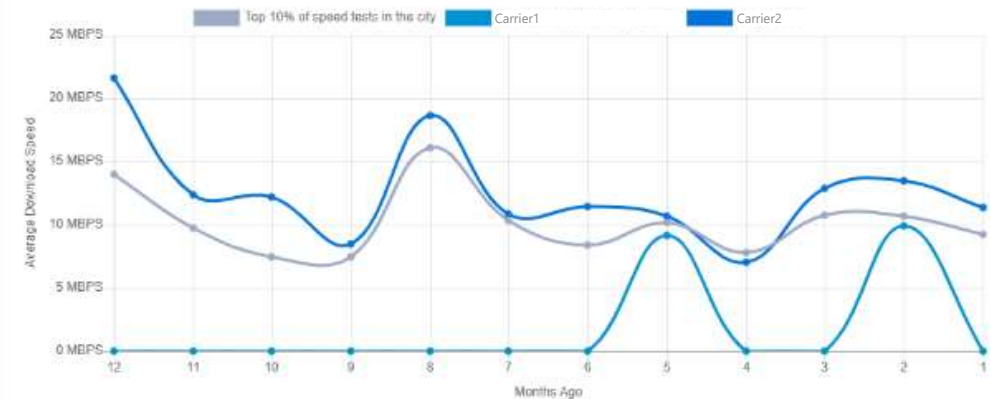
Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

7.7%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

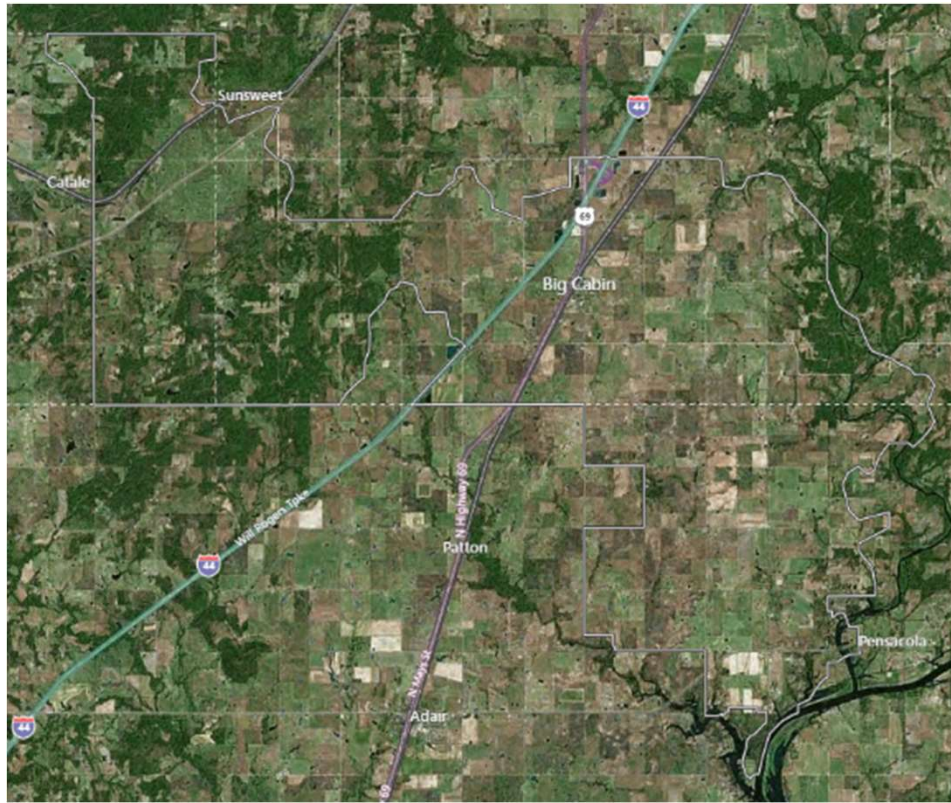
DOWNLOAD SPEEDS IN ROYAL



This analysis is based on 684 speed tests from IP-verified users who took speed tests from an IP address in Royal between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.

Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com

# Zip code: 74332 in Oklahoma



Estimated FCC broadband availability\*  
(2019 report)

99.9%

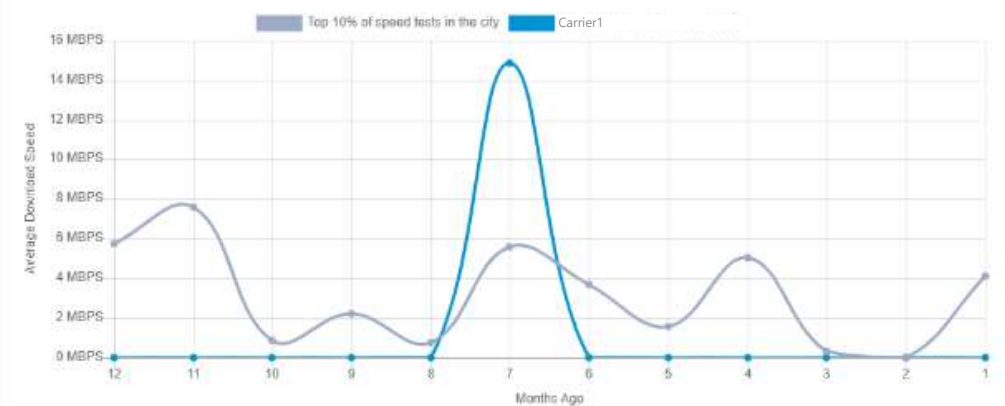
Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

0.6%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

DOWNLOAD SPEEDS IN BIG CABIN

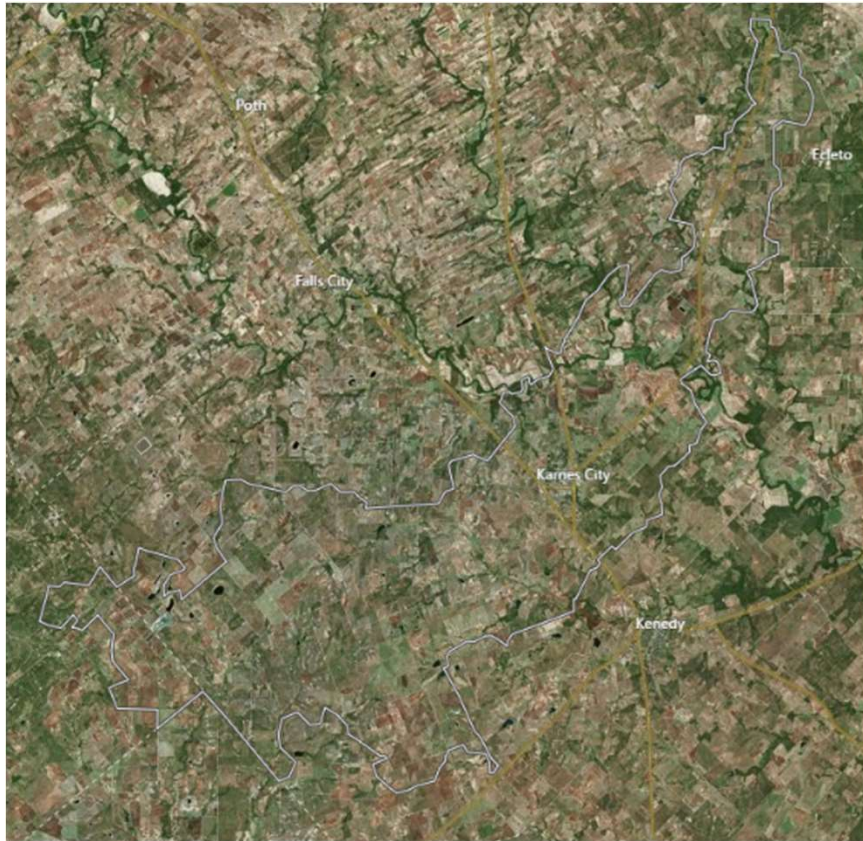


This analysis is based on 120 speed tests from IP verified users who took speed tests from an IP address in Big Cabin between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.

Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com



# Zip code: 78118 in Texas

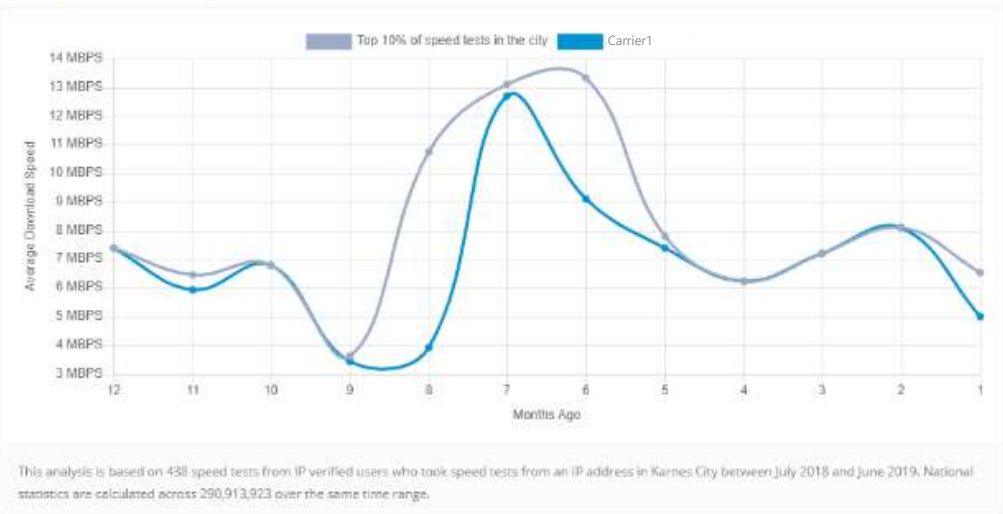


Estimated FCC broadband availability* (2019 report)	Estimated percent of people using the internet at broadband speeds using Microsoft data
100.0%	3.4%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

DOWNLOAD SPEEDS IN KARNES CITY



Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com



# Zip code: 78151 in Texas



Estimated FCC broadband availability\*  
(2019 report)

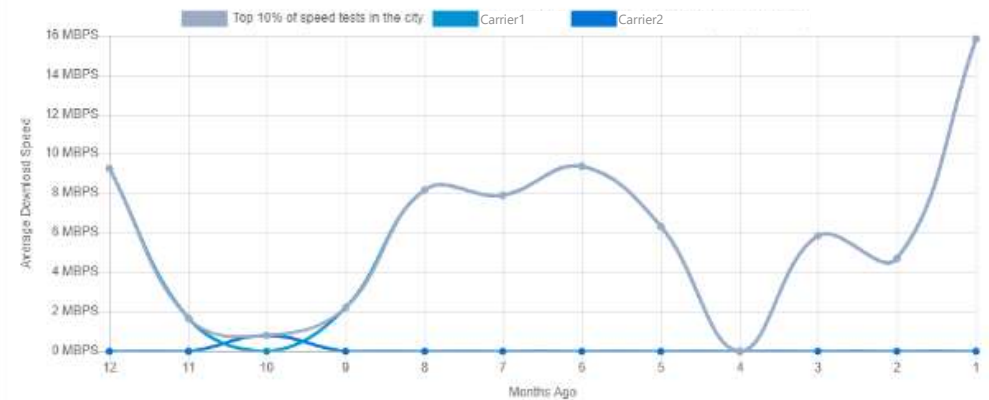
99.6%

Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

0.5%

\* zip codes may contain portions of multiple census tracts

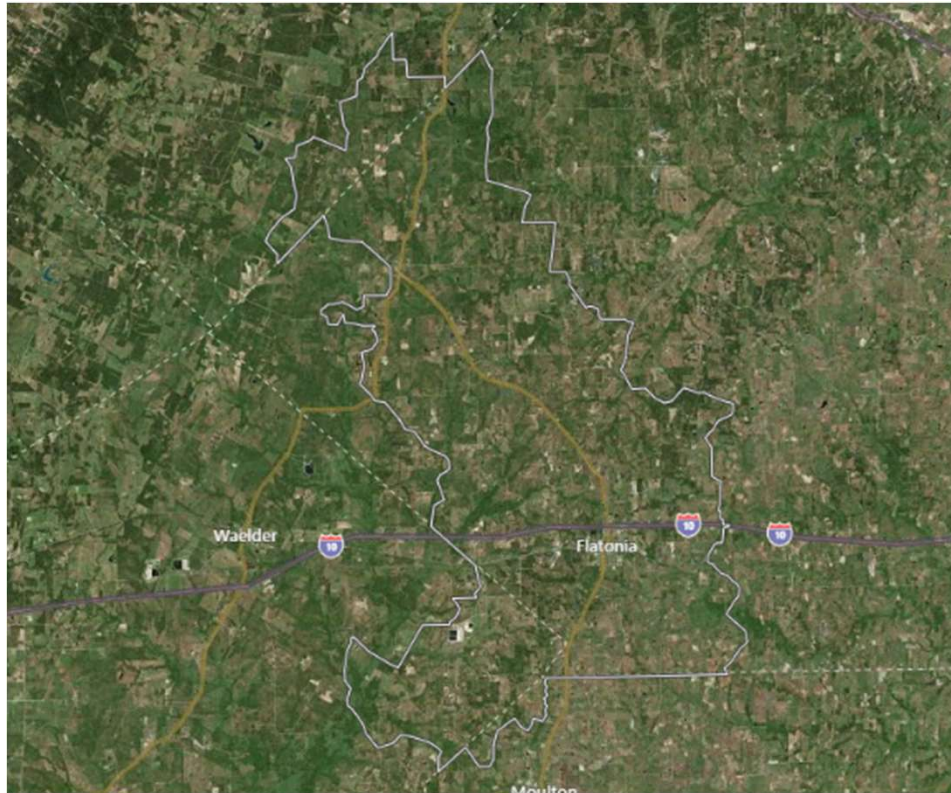
## BROADBANDNOW® DOWNLOAD SPEEDS IN RUNGE



This analysis is based on 119 speed tests from IP verified users who took speed tests from an IP address in Runge between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.

Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com

# Zip code: 78941 in Texas



Estimated FCC broadband availability\*  
(2019 report)

99.5%

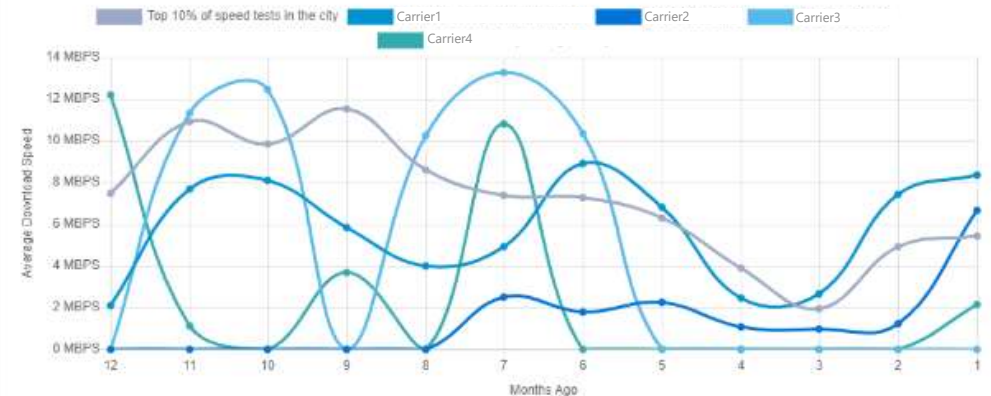
Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

2.8%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

DOWNLOAD SPEEDS IN FLATONIA



This analysis is based on 400 speed tests from IP verified users who took speed tests from an IP address in Flatonia between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.

Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com



# Zip code: 93602 in California



Estimated FCC broadband availability\*  
(2019 report)

93.7%

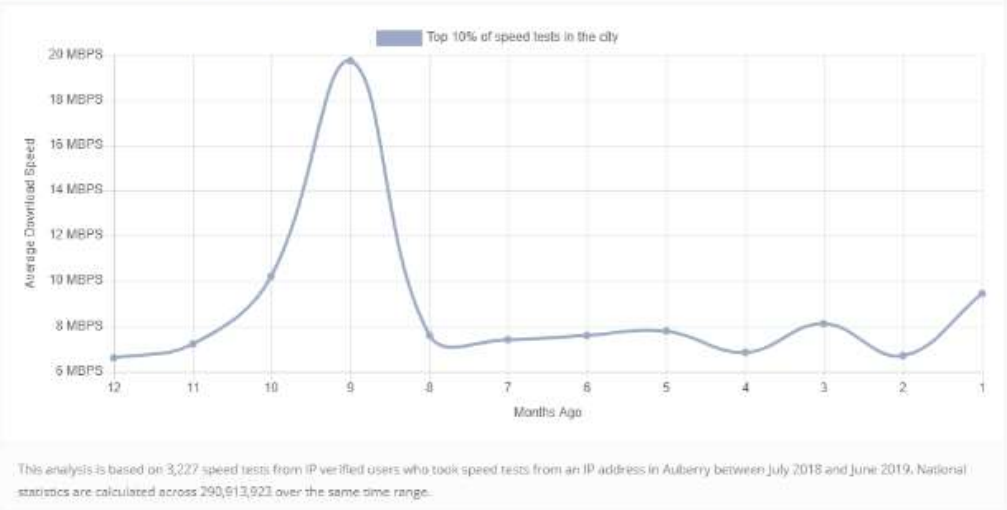
Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

8.5%

\* zip codes may contain portions of multiple census tracts

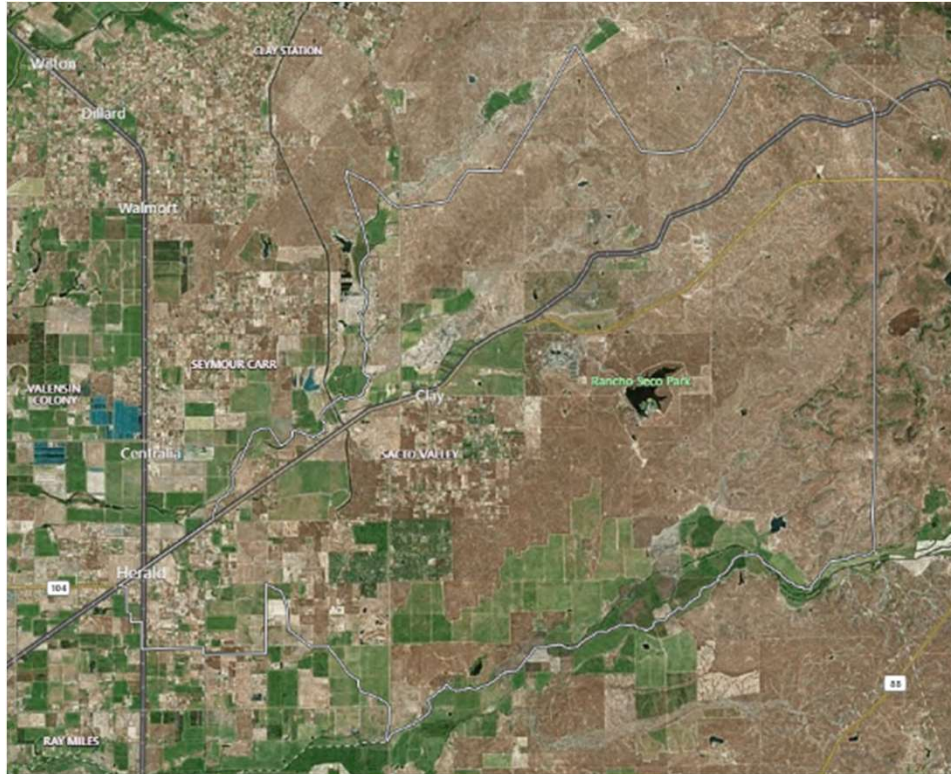
BROADBANDNOW®

DOWNLOAD SPEEDS IN AUBERRY



Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com

# Zip code: 95638 in California



Estimated FCC broadband availability\*  
(2019 report)

100.0%

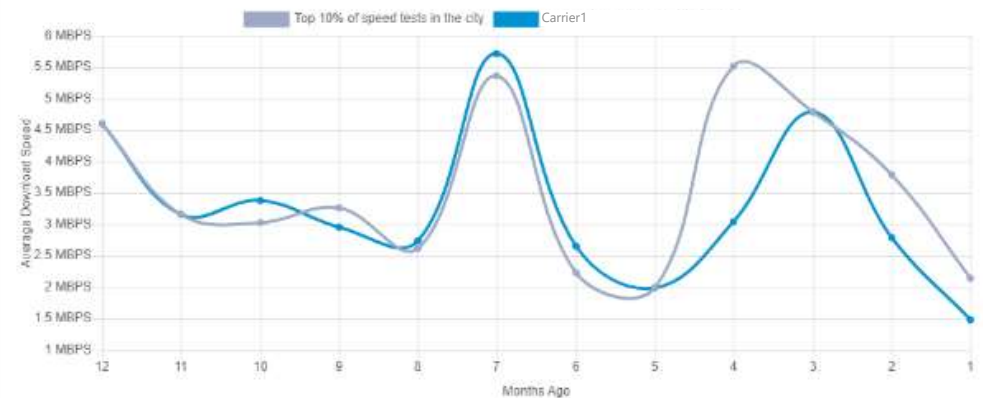
Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

2.2%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

DOWNLOAD SPEEDS IN HERALD

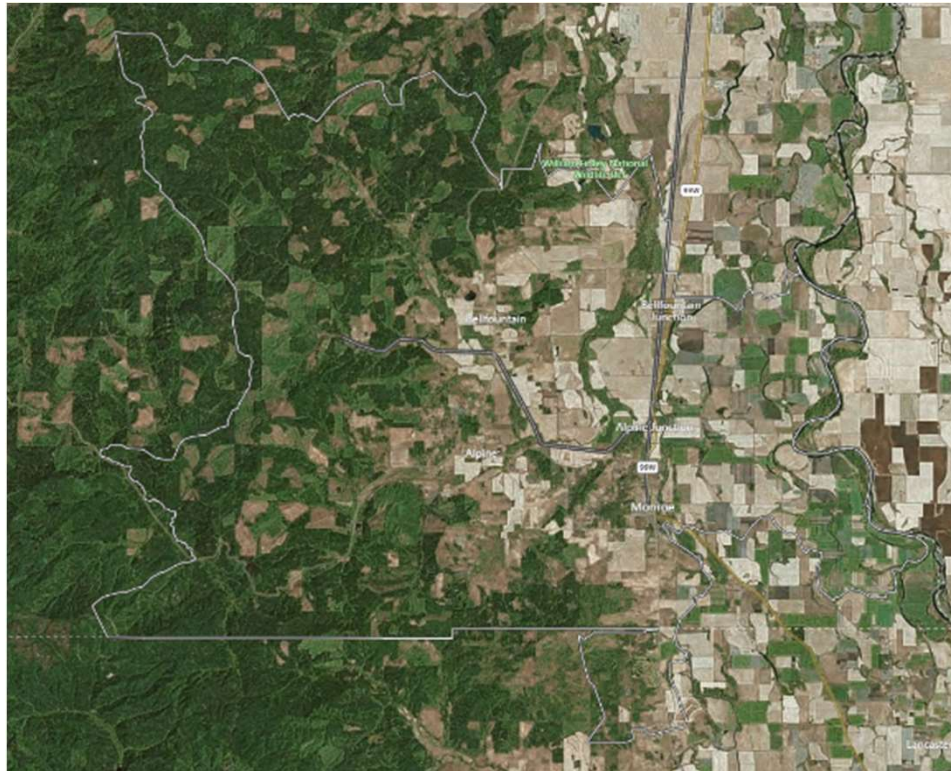


This analysis is based on 772 speed tests from IP-verified users who took speed tests from an IP address in Herald between July 2018 and June 2019; National statistics are calculated across 290,913,923 over the same time range.

Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com



# Zip code: 97456 in Oregon



Estimated FCC broadband availability\*  
(2019 report)

94.7%

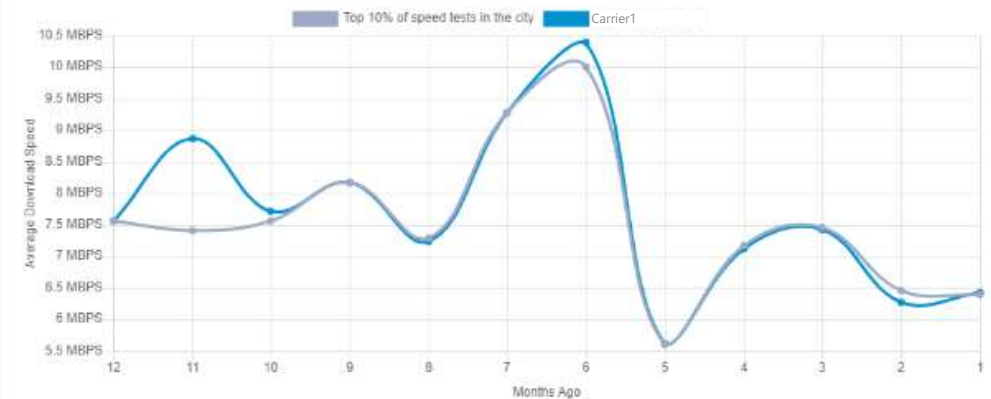
Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

7.6%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

DOWNLOAD SPEEDS IN MONROE



This analysis is based on 582 speed tests from IP-verified users who took speed tests from an IP address in Monroe between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.

Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com

# Zip code: 98855 in Washington



Estimated FCC broadband availability\*  
(2019 report)

97.9%

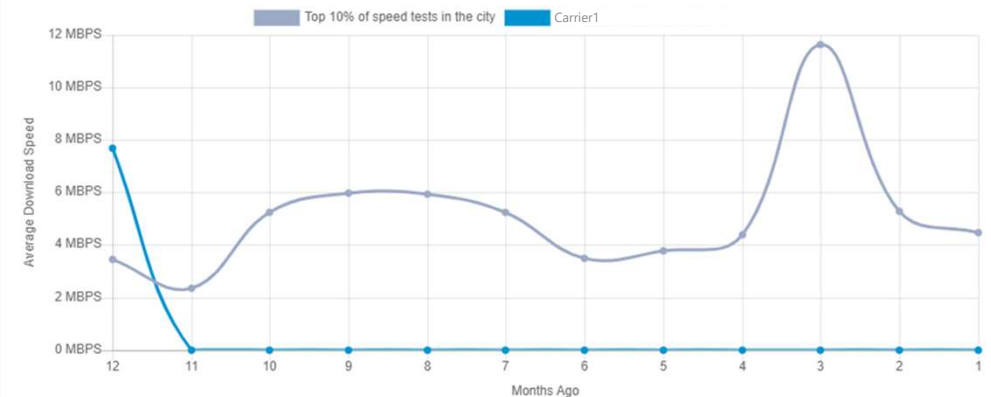
Estimated percent of people using the  
internet at broadband speeds using  
Microsoft data

7.4%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

DOWNLOAD SPEEDS IN TONASKET

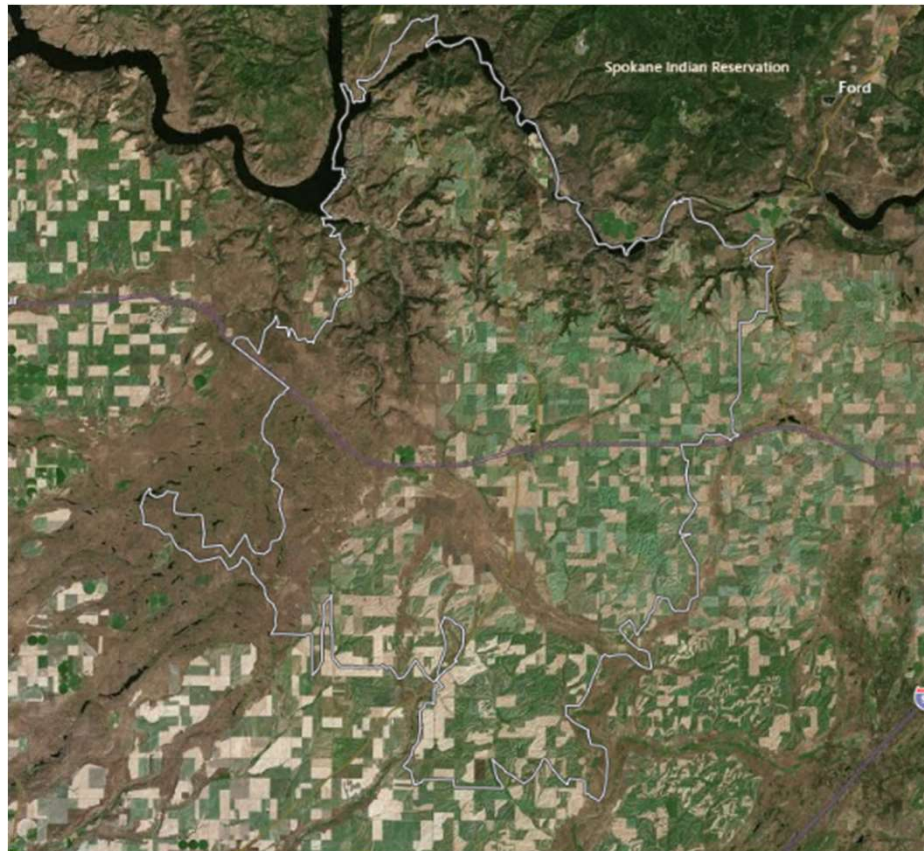


This analysis is based on 2,526 speed tests from IP verified users who took speed tests from an IP address in Tonasket between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.

Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com



# Zip code: 99122 in Washington



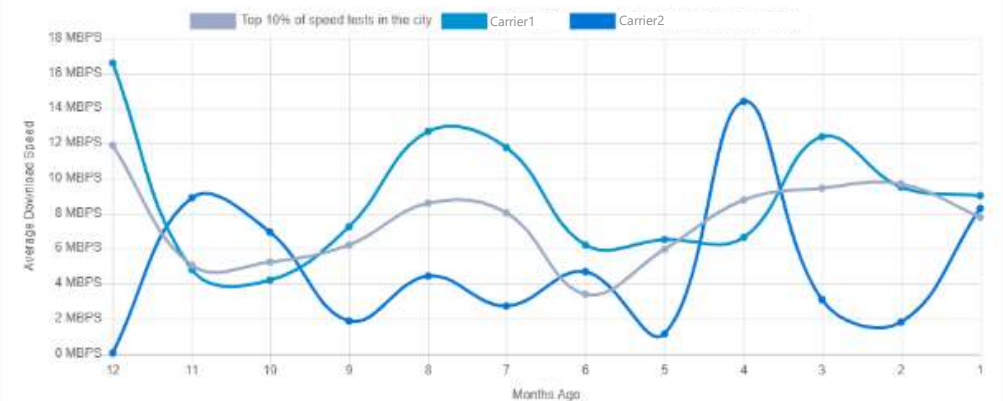
Source: FCC 2019 broadband report, Microsoft data, and BroadbandNow.com

Estimated FCC broadband availability* (2019 report)	Estimated percent of people using the internet at broadband speeds using Microsoft data
100.0%	4.0%

\* zip codes may contain portions of multiple census tracts

BROADBANDNOW®

## DOWNLOAD SPEEDS IN DAVENPORT



This analysis is based on 1,595 speed tests from IP verified users who took speed tests from an IP address in Davenport between July 2018 and June 2019. National statistics are calculated across 290,913,923 over the same time range.



**END**