Investigating **Variation in Broadband Access by Region Across the UK**

- Fardowsa Abdulkadir, Kim Clements, Isabelle Felsner, Yehjoon Kim, Lydia Maisander, Chinwendu Nwaozuzu -

PRESENTATION **OUTLINE**

- **INTRODUCTION**
- **DATA ANALYSIS**
- **CONCLUSION**
- **SWOT**

INTRODUCTION

Access to the internet is a vital part of modern life, highlighted during the pandemic.

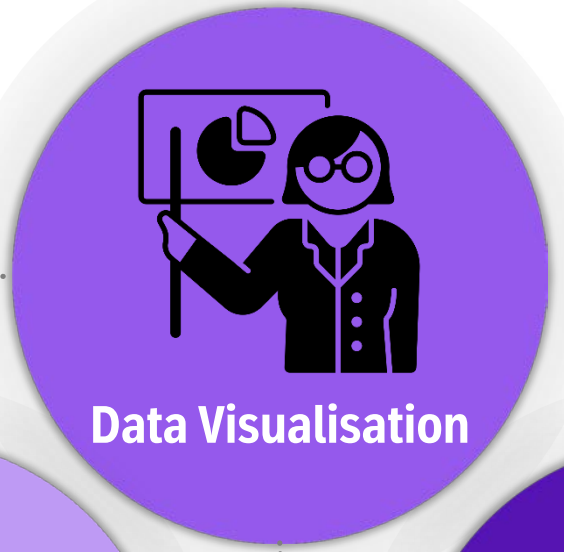
In recent years there has been heavy investment into faster fibre links across the country, however from our own personal experiences, we found that the UK's broadband infrastructure is not equal across all areas.

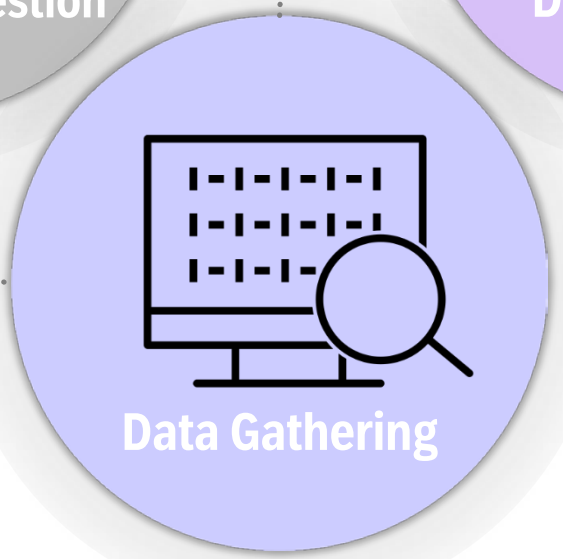
Our project aims to investigate and consider possible causes and potential solutions with regard to improving equality of access.

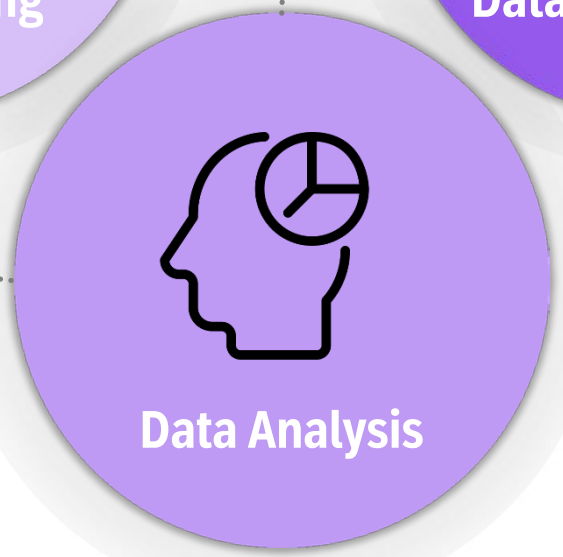
Variation in Broadband Access by Region Across the UK

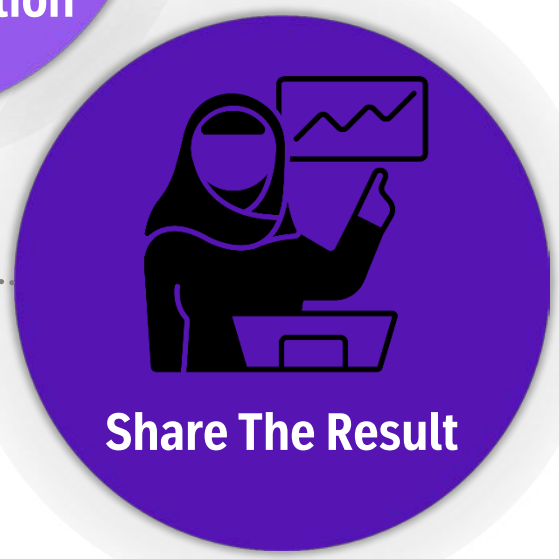
DATA ANALYSIS

Data Analysis Process









DATA ANALYSIS

Tools and Libraries

Library	Method	Section of Jupyter Notebooks (divided by questions)
Pandas	Dataframes, CSV and JSON files	All questions
Matplotlib	Data visualisation	All questions
NumPy	Machine learning formula	Question 2
Requests	API requests, using JSON	Question 2
Seaborn	Univariate, Bivariate and Multivariate plotting	All questions
Folium	Geospatial data visualisation	Question 1
SciKit-Learn	Supervised machine learning, predictive analysis models	Question 2

DATA ANALYSIS

Define Questions

Q1.

What is the difference of infrastructure between Rural and Urban areas?

Q2.

Can we identify trends over time?

Q3.

What is the nature of the relationship between internet and income at this moment in time?

DATA ANALYSIS

Data Gathering – Data Sources

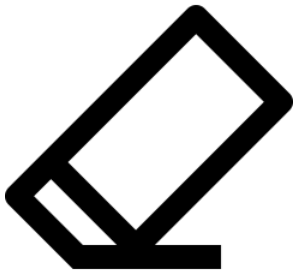
Income API

Rural/Urban Classification

Internet Coverage Data

DATA ANALYSIS

Data Cleaning



```
# Prepare income data for wide merge (wide dataframe)
income_wide = income_data.pivot(index=('Area', 'Region'), columns='Date', values='Observed Annual Income')
income = income_wide.rename_axis(None, axis=1).reset_index()
income
```

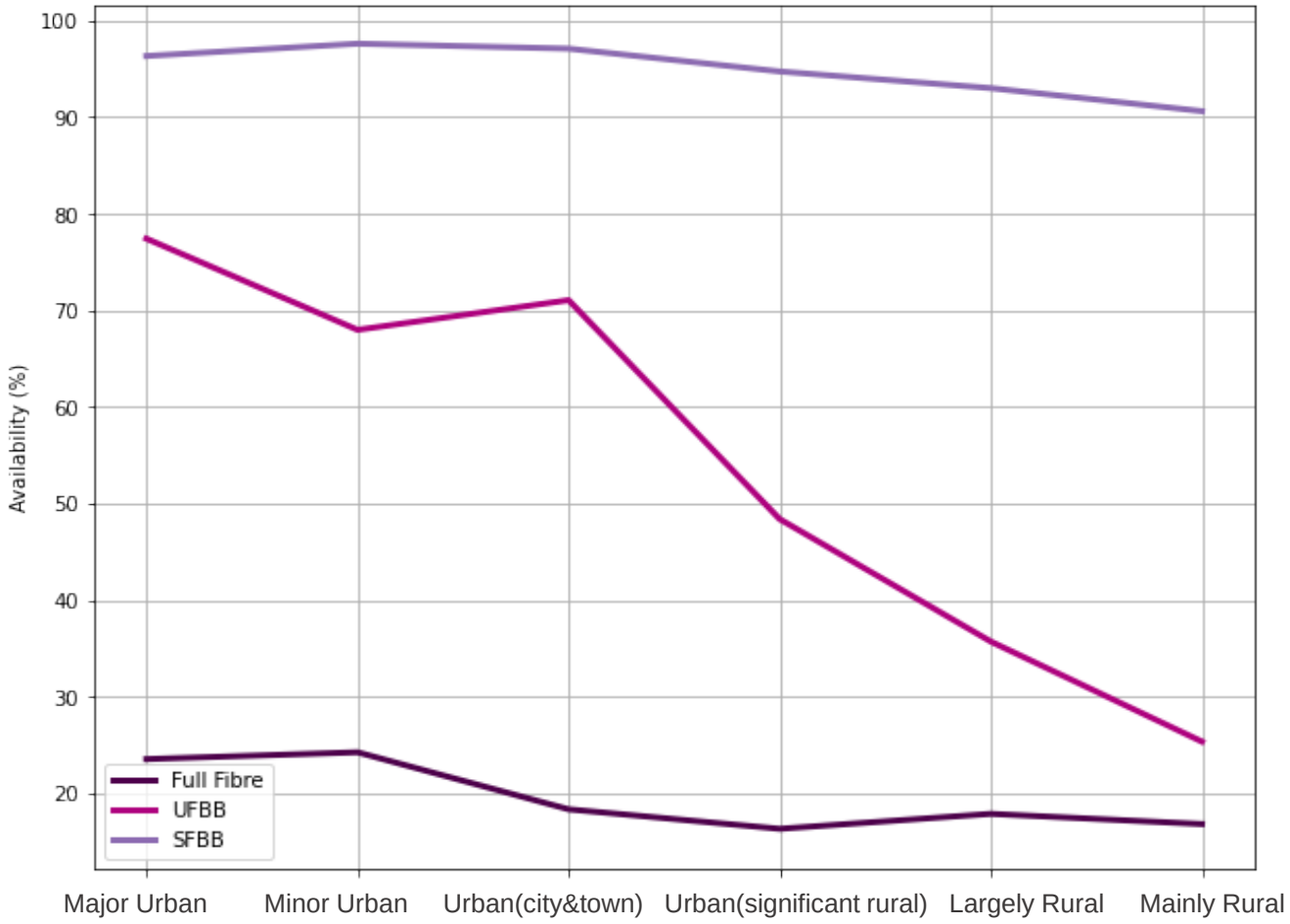
	Area	Region	2016	2017	2018	2019	2020	2021
0	ABERDEEN CITY	Scotland	29350.0	28500.0	29829.0	30426.0	30615.0	29949.0
1	ABERDEENSHIRE	Scotland	29887.0	29434.0	30877.0	32539.0	33970.0	32605.0
2	ADUR	South East	24645.0	24527.0	26771.0	27919.0	30250.0	NaN

```
nan_values = rural_urban_data[rural_urban_data.isna().any(axis=1)]
print (nan_values)
```

```
    laua Area Classification  Numerical classification
0    NaN  NaN             NaN                       NaN
327  NaN  NaN             NaN                       NaN
328  NaN  NaN             NaN                       NaN
```

DATA ANALYSIS

Data Visualisation - Q1. What is the difference of infrastructure between Rural and Urban areas?



Major Urban: Urban with Major Conurbation

Minor Urban: Urban with Minor Conurbation

Urban(city&town): Urban with City and Town

Urban(significant rural): Urban with Significant Rural

(rural including hub towns 26-49%)

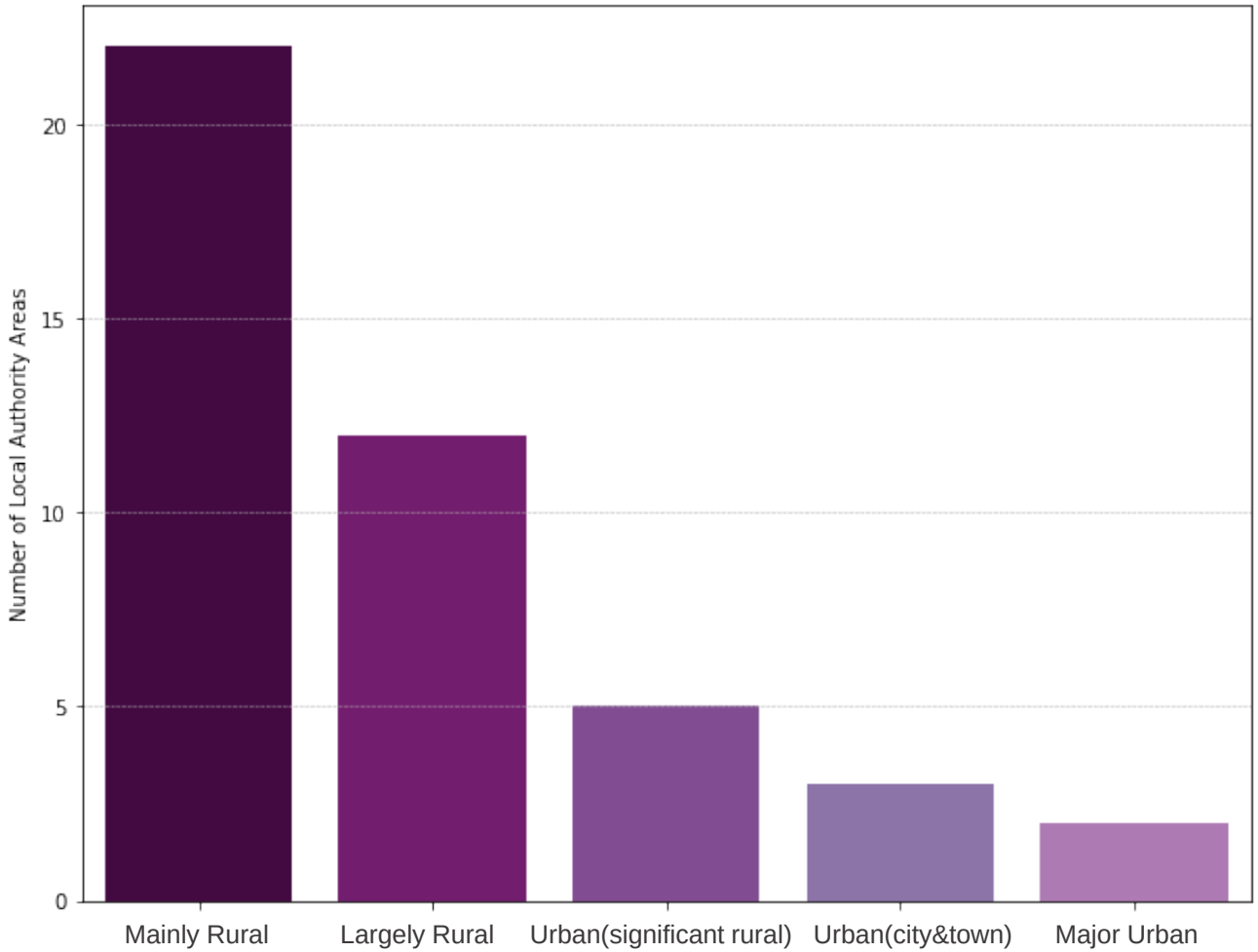
Largely Rural: Largely Rural (rural including hub towns 50-79%)

Mainly Rural: Mainly Rural (rural including hub towns >=80%)

Line plots showing the broadband availability of across different types of location (rural or urban).

DATA ANALYSIS

Data Visualisation - Q1. What is the difference of infrastructure between Rural and Urban areas?



Major Urban: Urban with Major Conurbation

Minor Urban: Urban with Minor Conurbation

Urban(city&town): Urban with City and Town

Urban(significant rural): Urban with Significant Rural

(rural including hub towns 26-49%)

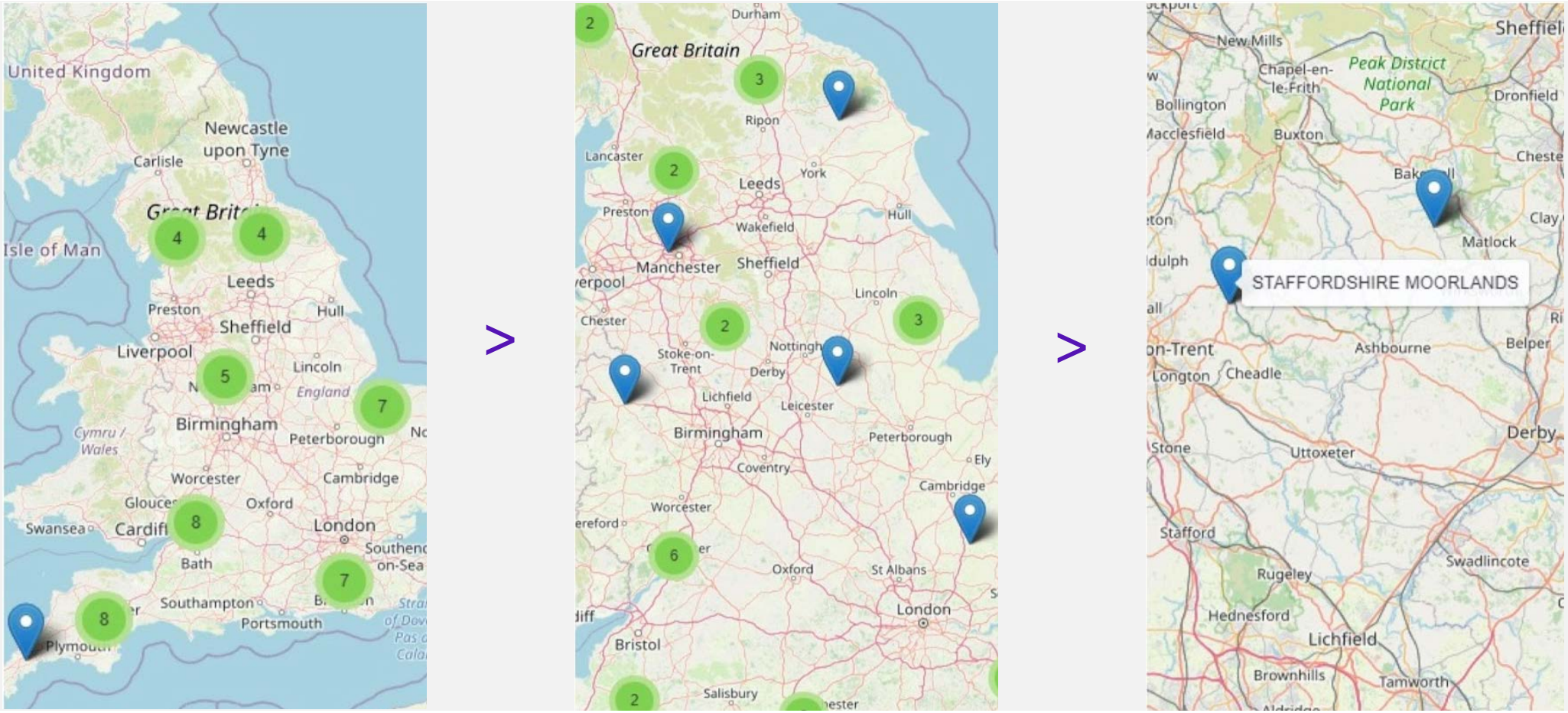
Largely Rural: Largely Rural (rural including hub towns 50-79%)

Mainly Rural: Mainly Rural (rural including hub towns >=80%)

This histogram shows the number of Local Authority Areas that do not meet minimum speed requirements of 8MBps for each Rural/Urban classification.

DATA ANALYSIS

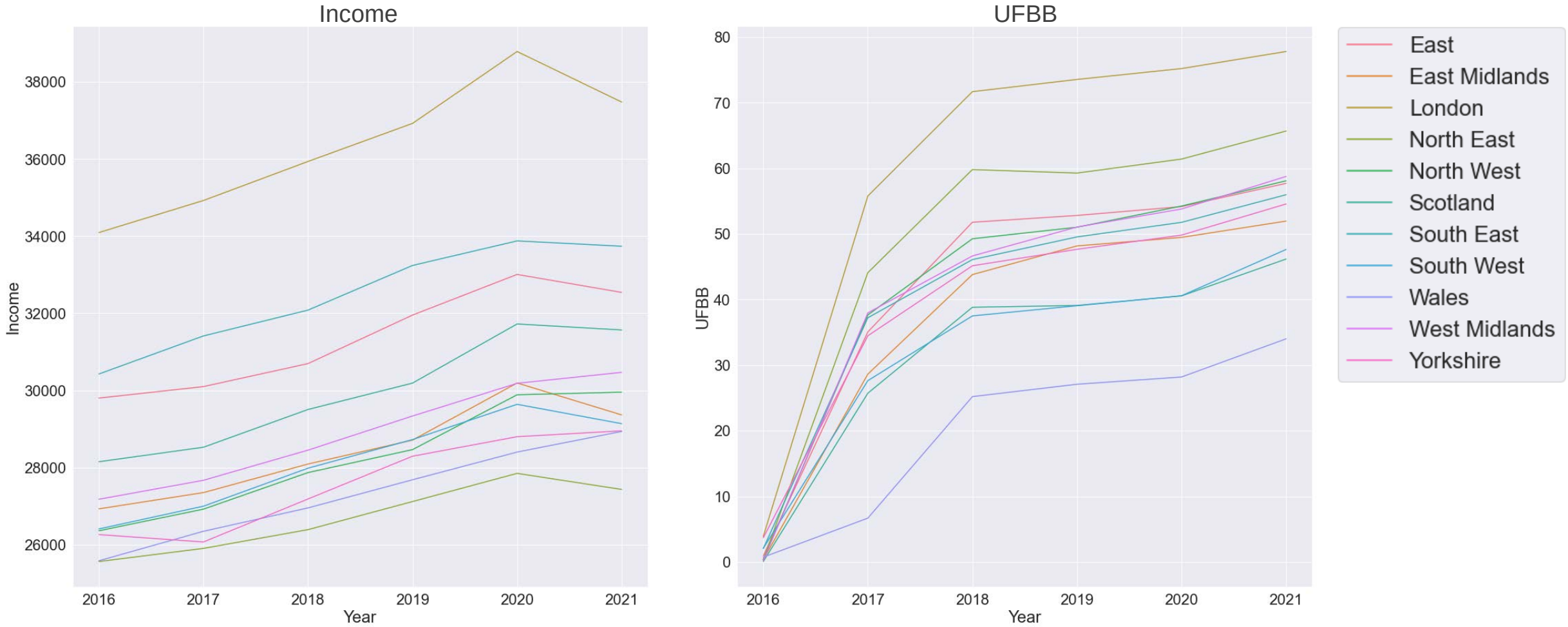
Data Visualisation - Q1. What is the difference of infrastructure between Rural and Urban areas?



The plotted map shows the Local Authority Areas where the minimum requirement does not meet 8MBps in a significant number of premises. This can be used by telecommunications companies to highlight areas which would benefit from infrastructure investment schemes that would improve speeds for users who are currently affected.

DATA ANALYSIS

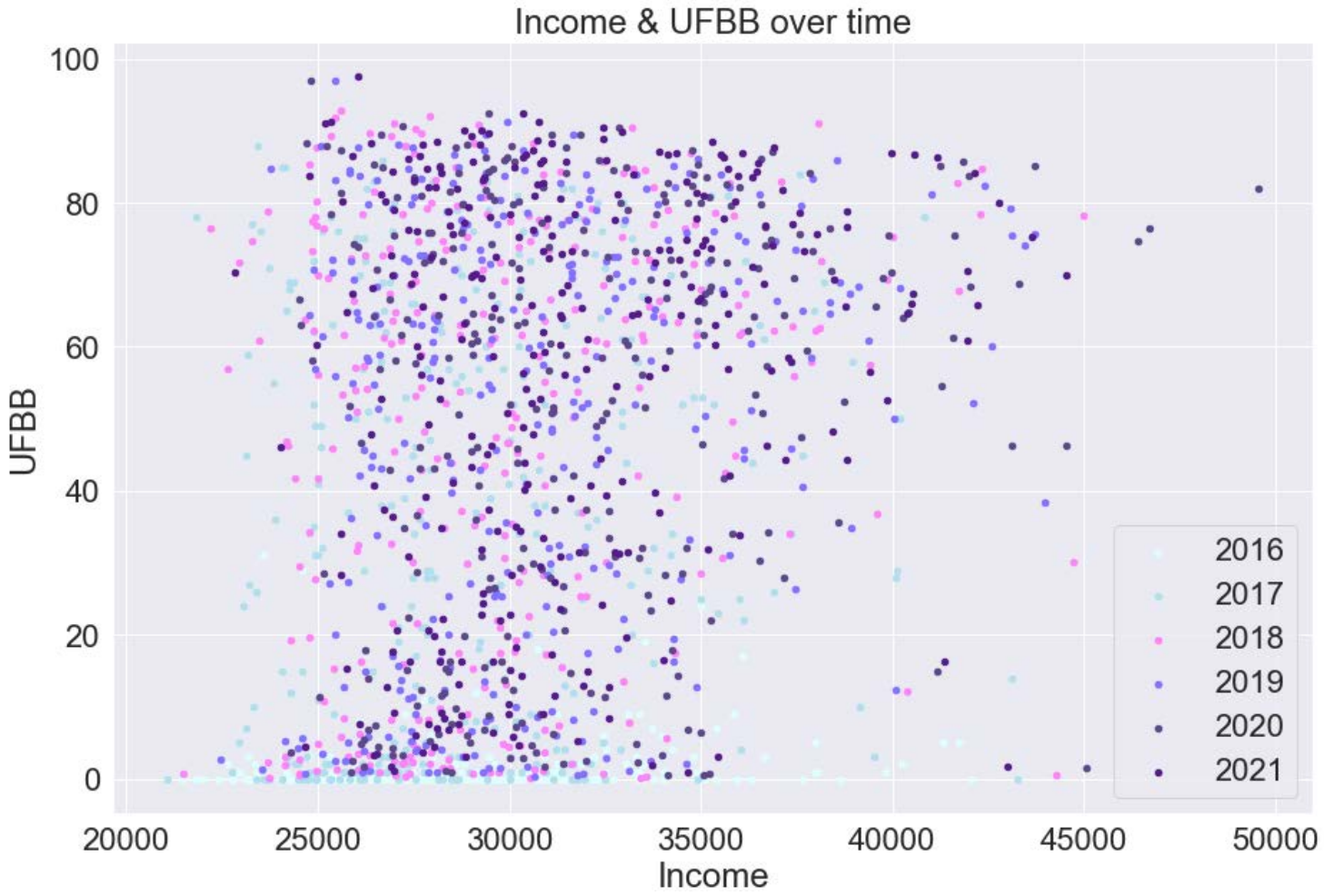
Data Visualisation - Q2. Can we identify trends over time?



Line plots showing the development of Income and UFBB over time, by region.

DATA ANALYSIS

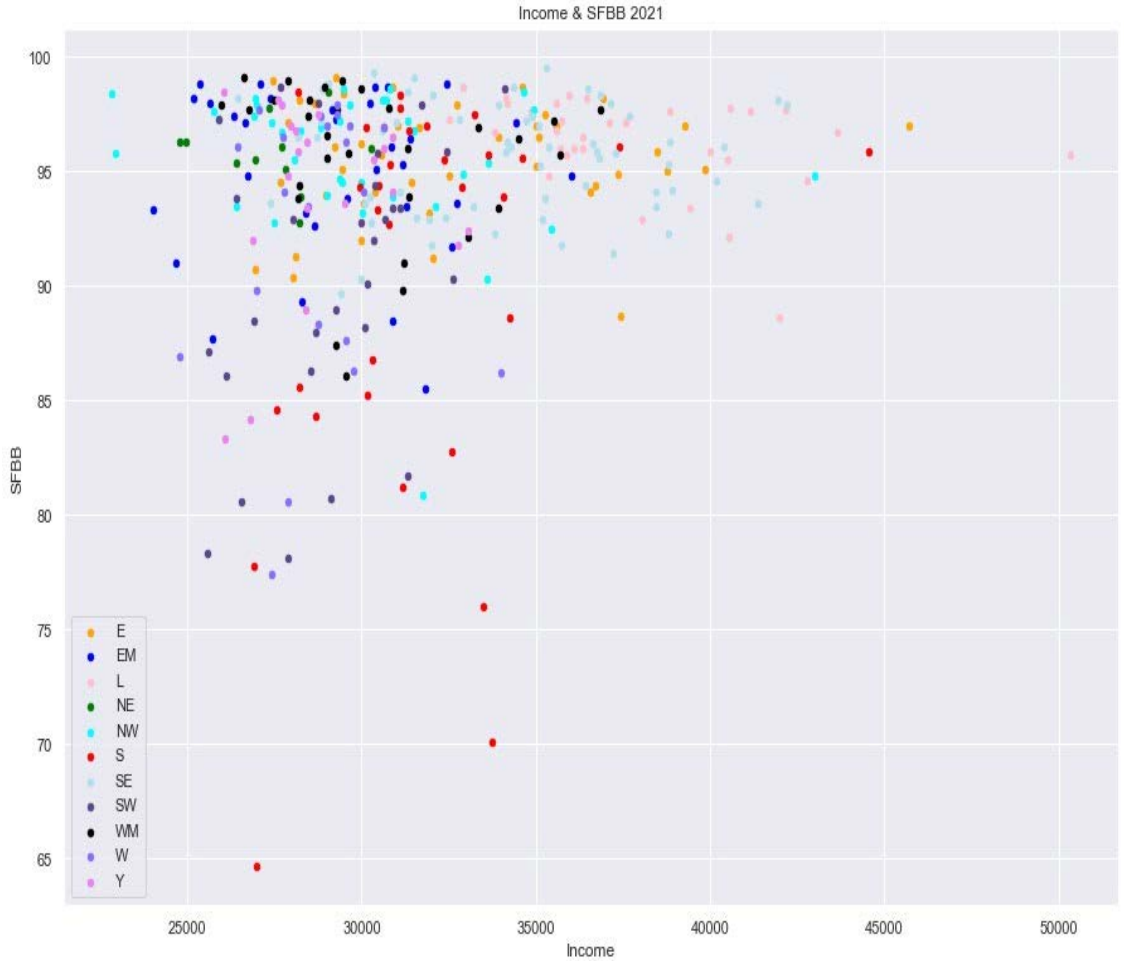
Data Visualisation - Q2. Can we identify trends over time?



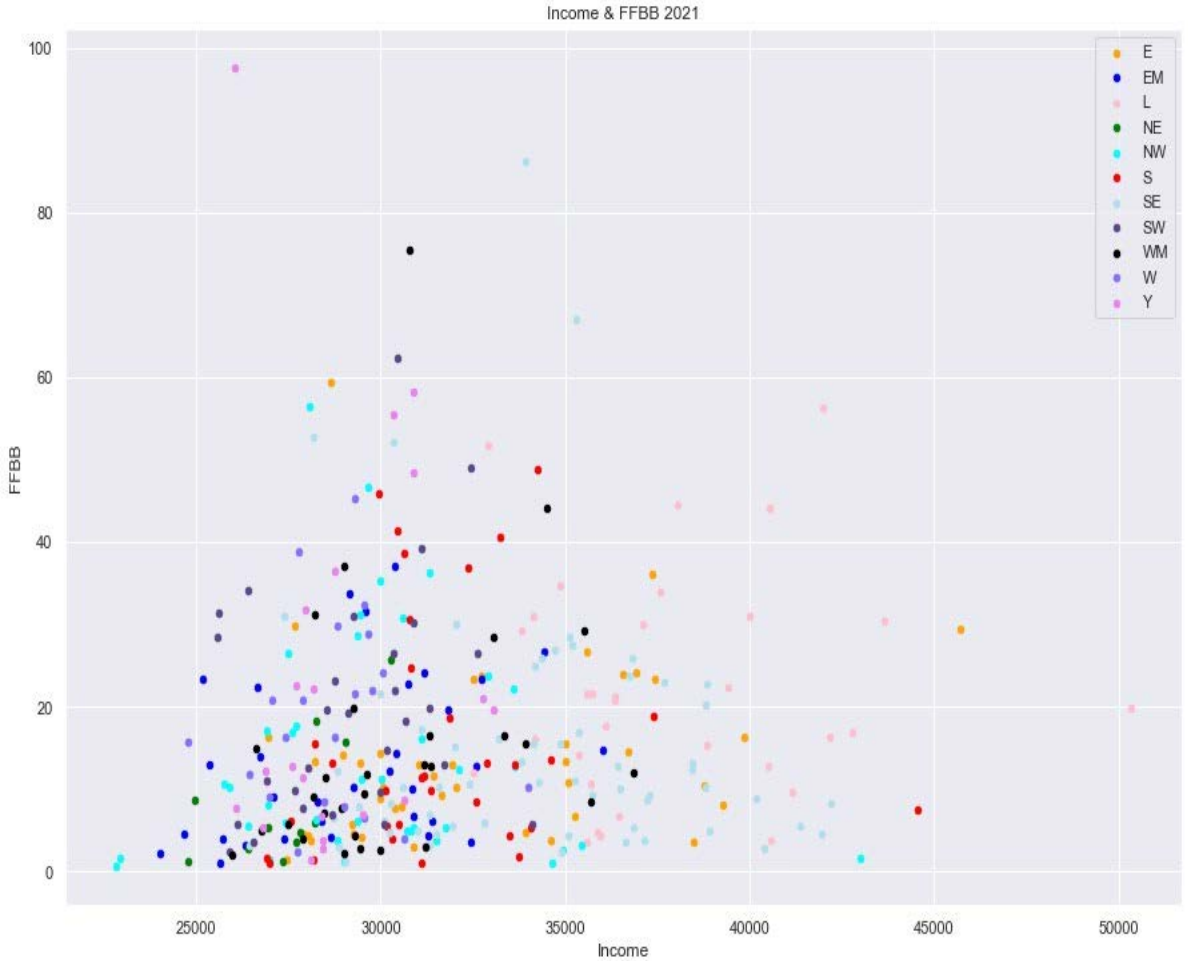
Scatterplot showing the relationship of Income and UFBB with time indicated by colour.

DATA ANALYSIS

Data Visualisation - Q3. What is the nature of the relationship between internet infrastructure and income at this moment in time?



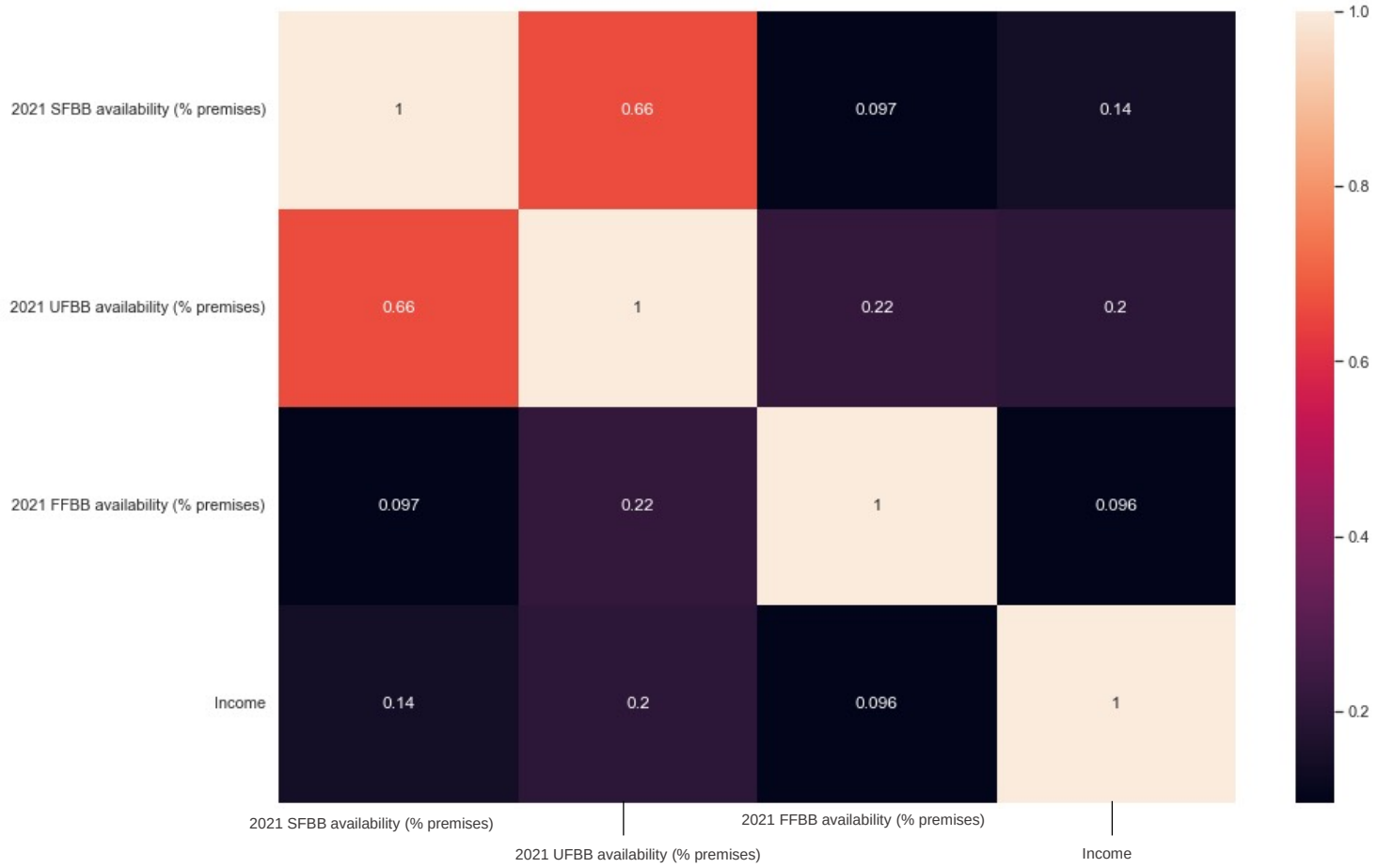
Scatterplot for SFBB vs Income



Scatter plot of Income and Full Fibre Broadband

DATA ANALYSIS

Data Visualisation - Q3. What is the nature of the relationship between internet infrastructure and income at this moment in time?



Correlation Matrix

Correlation Strength	Positive	Negative
Perfect	r = 0.9 to 1	r = -0.9 to -1
Strong	r = 0.5 to 0.9	r = -0.5 to -0.9
Weak	r = -0.1 to 0.5	r = -0.1 to -0.5
Uncorrelated	r = 0 to 0.1	r = 0 to -0.1

Note: any correlation above 0.3 and below -0.3 is considered significant

CONCLUSION

In 2021, most areas have a high coverage of SFBB, regardless of whether it is rural or urban. Additionally, no patterning with the income of an area can be identified. There were some trends visible for certain regions, however the correlation coefficient was non-significant.

In 2021, very few areas have FFBB coverage, regardless of whether it is rural or urban. Again, income did not have an influence.

In 2021, UFBB showed a clear difference in availability between rural and urban areas, with urban areas having a much higher percentage of premises who had UFBB available to them. Income was not a predictor of UFBB coverage.

Over time, the UFBB coverage in an area could not be predicted by income. Both income and UFBB have been continuously trending upwards since 2016.

A clear trend between rural and urban areas can be seen, when looking at internet speeds. A significant number of rural areas only have access to inadequate speeds.

SWOT ANALYSIS

STRENGTHS

- Effective data analysis
- API integration
- Teamwork : efficient communication
- Good use of python libraries and packages :
  using complex libraries, exploring new data visualization techniques - e.g. folium
- Transferable skill set across team from diverse field (Technical & Non-Technical)

WEAKNESSES

- SQL database not used
- Scheduling issues for sprint meetings/team members not always available
- Large team split into sub groups
  - KISS methodology not utilised

OPPORTUNITIES

- Teams varied background :
  maximized prior experiences when analyzing data
  - e.g. predictive analysis skills
- Online resources utilized to improve understanding of key topics

THREATS

- API key permissions not provided on time
- Availability of data sets
- Timing for project - holidays took precedence
- Python knowledge
  - data visualisation tools not fully understood

QUESTIONS

WA MAHADSANID

THANK YOU

DANKE SCHÖN

감사합니다.

DAALU