**Exploratory Data Analysis of Claimant Count in the UK**

**Introduction**

The Claimant count is the number of people claiming benefits primarily for the reason of being unemployed. This is measured by combining the number of people claiming Jobseeker’s Allowance, National Insurance credits and receiving Universal Credits [1]. The claimant count and unemployment are both important measures of labour force in the UK. However, the claimant count is mostly derived from Department for Work and Pensions (DWP) administrative systems, while unemployment based on individual’s self-declaration as being out of work, but currently and actively seeking to work in the Labour Force Survey (LFS).

The increased number of claimant count, especially during pandemic has long lasting impacts on individuals, economy, and society. According to IFS (Institute of Fiscal Studies), during the pandemic, JSA (Job Seekers Allowance) volumes began to rise, and achieved its highest level in August 2020 since August 2018. Latest JSA figures to February 2021 show that there were 260,000 people claiming Jobseeker’s Allowance (JSA), which is still an increase of 96,000 (57%) on the pre-pandemic figure of 170,000 at February 2020[2].

According to an article in BBC, the government spend about £100bn a year on the working age people benefits than it does on educational or national defence and policing [3]. It showed the importance of spare or out of work labour in the nation’s broader economy and well-being of the society. It has long lasting impact on individuals’ capability, and physical and mental health.

The government introduced furlough scheme during pandemic to protect the economy and to maintain the unemployment rate during the Covid-19. Reading these articles on increased amount of government spending on benefits has led me to explore the claimants count on local authorities’ (counties) level.

There is plenty research on unemployment and its effect on the economy and society, but there is hardly any research on claimant count. The researcher may use both interchangeably to review the economy outlook on the national level, there is no research on local level regarding the claimants count and its effect on the local society or residents.

The aim of this exploratory research is to find the trends of the claimants over the last decade and if there are any underlying factors of high claimants in certain counties of UK. I will also analyse unemployment data to observe any similarities and differences between the claimant’s data.

My objective is to gather the local authorities (counties) data on claimant count, categorised the data into 4 age groups, filter the data with any mislead or incorrect numbers, tabulate the data, then visualise the data, and find out the highest claimant’s count.

**Data**

**Data source and description**

The data for this project obtained from the Nomis website. Nomis is a service provided by [Office for National Statistics](https://www.ons.gov.uk/) (ONS), the UK’s largest independent producer of official statistics. The data has taken from 2013 to 2021 as Claimants count introduced in 2013 to replace Job seekers allowance (JSA). Currently the Claimant Count is a composite of the number of people claiming Jobseeker’s Allowance (JSA) and those claiming Universal Credit (UC) who are required to seek work to qualify for their benefits. Combining these two identifies all the people claiming benefit principally for the reason of being unemployed [4].

I have taken the data from Nomis [5] and construct my query on the Claimants count geographically by local authorities (as of April2019), select the dates from May 2022 to 2013, and categorise the data into 4 age groups: age over 16 years (working population), 16-24 years age group, 25-49 years, and last over 50 years of age.

I have taken unemployment data from Nomis as well from 2013 to 2021 as 2022 is still not published. I have taken this data to compare unemployment rate with the claimant rate and see if there are any similarities. I also analyse the working age population against the unemployment rate and with the unemployed working age population.

**Limitations of Data**

The claimant count and unemployment both have large degree of similarities between them. However, both have few limitations in terms of reporting the data. These are:

1. People with part time job can claim benefits if their earnings are very low. However, they would not appear in LFS measure of unemployment.
2. Those in full time education while searching for work will not appear under claimant count as they are not eligible for the benefits, which predominantly affect 16-21 years old. They will appear in unemployment data.
3. Those over or around state pension age who are looking for work will not be eligible for unemployment benefits and not report under claimant count. They will appear in the LFS measure of unemployment.
4. The people who are looking for part time work can not claim unemployment benefits and therefore will not be reported under claimant count but under unemployment. The affect of this claimant count can be seen particularly in women with school age children.

**Ethical Consideration of the Data**

The data is easily publicly accessible, and it has featured different datasets (Excel, XML, Csv and so on). The download has a limit of1000000 cells in a single download.

**Importing the Data**

I have imported the claimant’s data in CSV format as ‘claimants count\_local authority\_nomis.csv’.

I have also imported the unemployment data from Annual Population Survey in CSV format as well.

**Data Cleaning and Processing**

There are 217 counties or local authorities in UK and my interest is to select ten local counties with largest claimant counts during 2013-2022. I process ten largest claimants count because most of the claimant’s data is fairly condensed apart from ten local authorities. The ten local counties will be enough to graphically represent the data.

1. **Total Claimant Count of Age Above 16 Years**

First, I import the total claimant count of age above 16 years by using the function of skiprows and nrows into the table form through pandas and data frame.

Graphical user interface, table

Description automatically generated with medium confidence

Graphical user interface, application, table

Description automatically generated

* 1. **Distribution of All Claimant Count:**

Use histograms to visualise the distribution of the data. It can be clearly seen most of the claimants fall within the range of 10,000 between 2013-2019. The number of claimants increased to 20,000 from 2020 to 2022. The exception year was 2016, most of the local authorities claims fall within 5000 claimants. There is significant difference between the highest claimants in local authority to the second highest claimants in local county.

Chart, box and whisker chart

Description automatically generated

**Mean Value:**

Take out the mean value of each year to see if there are any large differences in values. Use numeric only function to get only numeric values.

Graphical user interface, text

Description automatically generated

It is quite evident that during the pandemic (2020-2021) the claimants count mean is nearly double as compared to pre-pandemic years (2014-2019) with only exception in 2013.

* 1. **Tabulating the Top 10 Highest Claimants counties of all Age 16+:**

Graphical user interface, application

Description automatically generated

**Graphical user interface

Description automatically generated with medium confidence**

**Graphical user interface, application

Description automatically generated**

**Graphical user interface, application

Description automatically generated**

**Graphical user interface, table

Description automatically generated**

**Table

Description automatically generated**

**Graphical user interface, application

Description automatically generated**

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

1. **Claimants as a Proportion of Residents Aged 16-64 years:**

I have loaded, filtered, and cleaned the selected data from my csv file using skiprow, nrows, and na\_values function. The proportion/rates for local authorities from 2020 onwards are calculated using the mid-2020 resident population aged 16-64. It will help me to analyse the claimant count against the working age population.

Graphical user interface, application

Description automatically generated with medium confidence

A picture containing text

Description automatically generated

**2.1 Tabulating the 10 Highest Total Claimant counties as Proportion of Residents Aged 16-64 Years:**

I find the 10 highest claimant counties as proportion of residents aged 16-64 years for each year by using the nlargest function.

Graphical user interface, text, application

Description automatically generated

Graphical user interface

Description automatically generated with low confidence

A picture containing graphical user interface

Description automatically generated

A picture containing text

Description automatically generated

Graphical user interface, application

Description automatically generated with medium confidence

Graphical user interface

Description automatically generated with low confidence

Graphical user interface, application

Description automatically generated with medium confidence

Graphical user interface, text

Description automatically generated with medium confidence

A picture containing graphical user interface

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence

1. **Claimants Count of Aged 16-24 Years:**

Import the data of aged 16-24 years to analyse the claimants count within this age group.

Table

Description automatically generated with medium confidence

**Graphical user interface, application

Description automatically generated**

* **3.1 The Distribution of Claimant Count of Aged 16-24 Years:**

**Chart

Description automatically generated**

Most of the local counties before pandemic fall within the range of 5-2000 claimant numbers with the exception in 2013. The claimants number reached to 5000 during pandemic. The number has been increased to nearly double as it has observed in the total claimant’s count.

* **3.2 Mean Value of the Claimant Counts of Aged 16-24 years:**

The mean of the claimant counts before pandemic apart from 2013 is under 1000. However, it reached to 2000 during 2020 and 2021. It showed the claimants number has same effect during pandemic as with total claimants’ count.

Graphical user interface, text

Description automatically generated

* **3.3 Tabulating the 10 Highest Total Claimant counties of Aged 16-24 Years:**

**Graphical user interface, application

Description automatically generated**

**Graphical user interface, application

Description automatically generated**

**Graphical user interface, application

Description automatically generated**

**Graphical user interface, application

Description automatically generated**

**Graphical user interface

Description automatically generated**

**Graphical user interface

Description automatically generated**

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

1. **Claimant Counts of Aged 25-49 Years:**

Select the data for this age group from 'claimants count\_local authority\_nomis.csv' file using the same function (skiprows and nrows) and clean the data from any non-numeric value by na\_values.

Graphical user interface, application, table

Description automatically generated

Graphical user interface, table

Description automatically generated with medium confidence

**4.1 The Distribution of Claimant Count of Aged 25-49 Years:**

I create the histograms to visualise the distribution of claimant count of aged 25-49 years.

Chart, box and whisker chart

Description automatically generated

The claimant number in this group of age shows similar pattern as in previous age group (16-24 years), most of the local authorities fall within 5000 claimants except for 2013. The figure has increased to 10,000 during the pandemic 2020 and 2021.

**Mean Values of Claimants of Aged 25-49 Years**

Text

Description automatically generated

* 1. **Tabulating the Ten Local Counties with Highest Claimant Counts of Aged 25-49 Years:**

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

1. **Claimant Counts of Aged Over 50 Years:**

Select the data for this age group from 'claimants count\_local authority\_nomis.csv' file using the same function (skiprows and nrows) and clean the data from any non-numeric value by na\_values.

Graphical user interface, table

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence

* **5.1 The Distribution of Claimant Count of Aged Over 50 Years:**

I plot the histogram to visualise the distribution of the data. The distribution of the data showed most of the local authorities’ claimant fall within 2000. However, it increased to 5000 in 2020 and 2021 during the pandemic.

Chart

Description automatically generated

**Mean Vaues of Claimants of Aged over 50 Years:**

**Text

Description automatically generated**

* 1. **Tabulating the Ten Local Counties with Highest Claimant Counts of Age Over 50 Years:**

**Graphical user interface, application

Description automatically generated**

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface

Description automatically generated

**Unemployment Data**

I have constructed my query from Annual Population Survey data and import it to csv format. Then I load the data into dataframe. It will help me to compare the unemployment rate with the claimant rate and see if there are any similarities. The denominator column reflects the working age population and numerator unemployed working age population.

Table

Description automatically generated

Table

Description automatically generated

**Tabulating Ten Local Authorities with Highest Unemployment rate:**

Find the ten largest unemployment rates from 2013-2021 by using nlargest function and remove the dtype error by coercing the data.

Graphical user interface, table

Description automatically generated

Table

Description automatically generated with low confidence

Graphical user interface, application

Description automatically generated

Table

Description automatically generated with medium confidence

Graphical user interface

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence

A picture containing graphical user interface

Description automatically generated

Table

Description automatically generated with medium confidence

A picture containing table

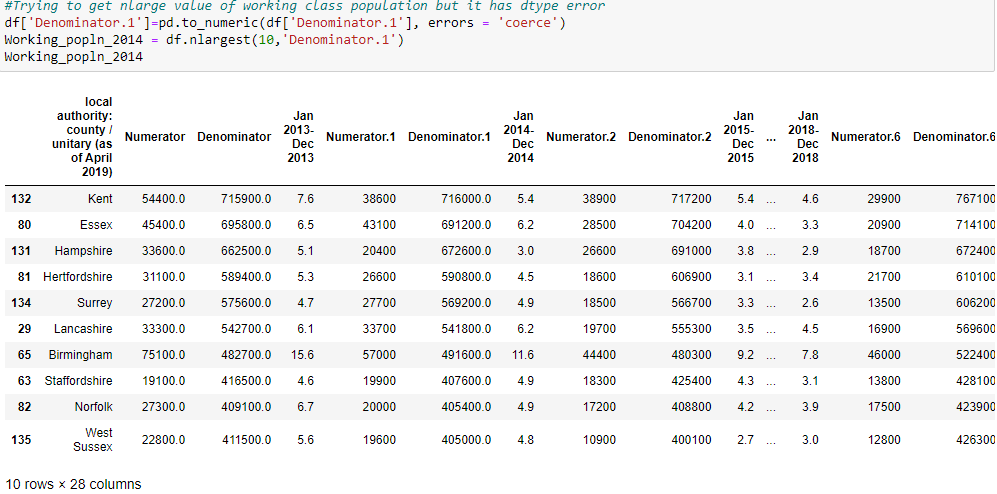
Description automatically generated

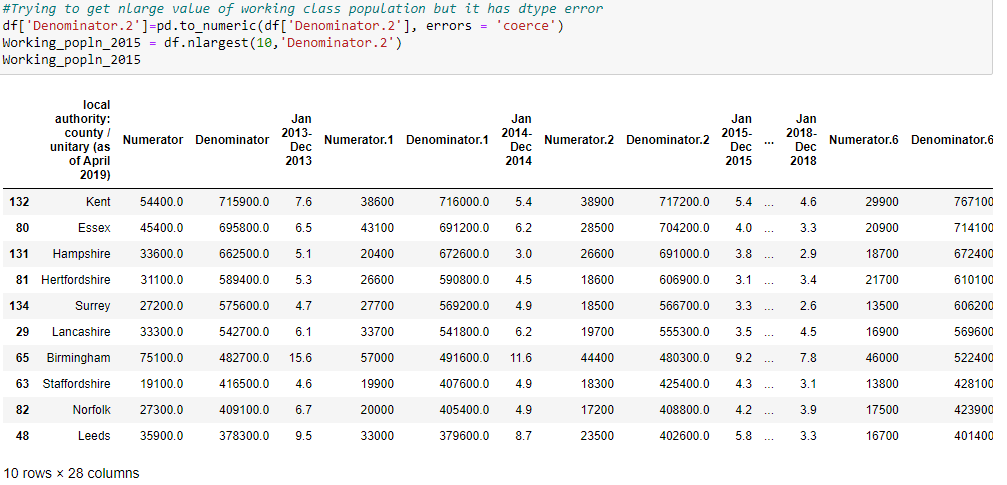
**Tabulating Ten Local Counties with Largest Working Population**

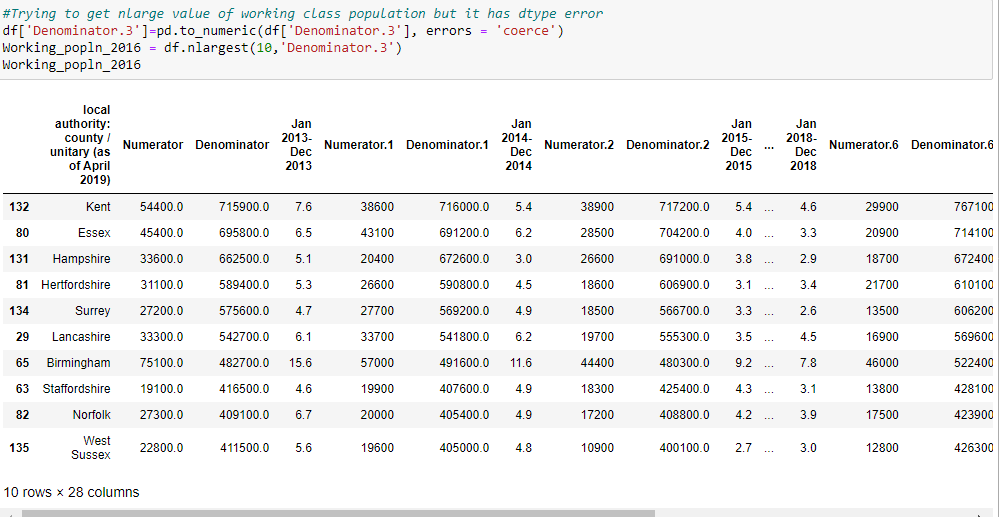
The denominator column reflects the working age population. The unemployment rate is related to working age population of local authorities and if the working age population is small then it can give high or low unemployment rate. So, I am finding out the ten largest working age population and it will give better understanding to analyse the unemployment rate and claimants.

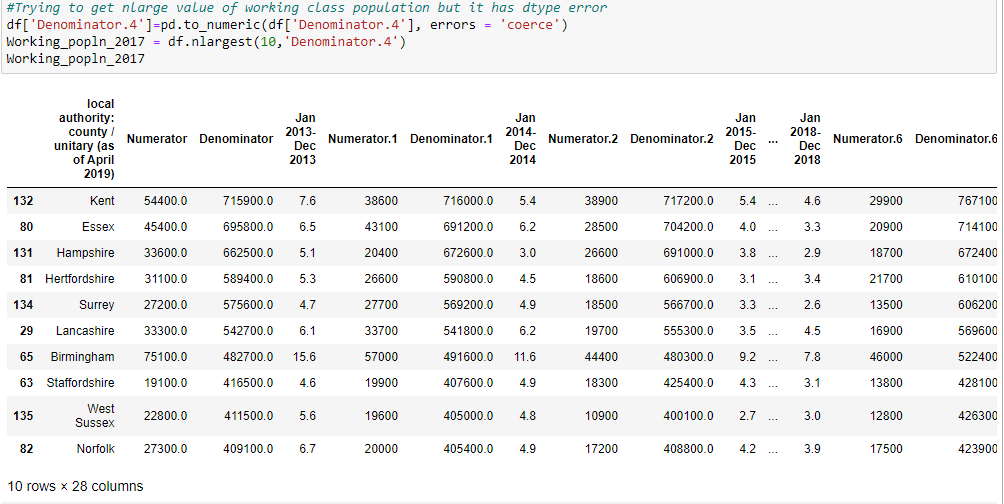
Graphical user interface, application

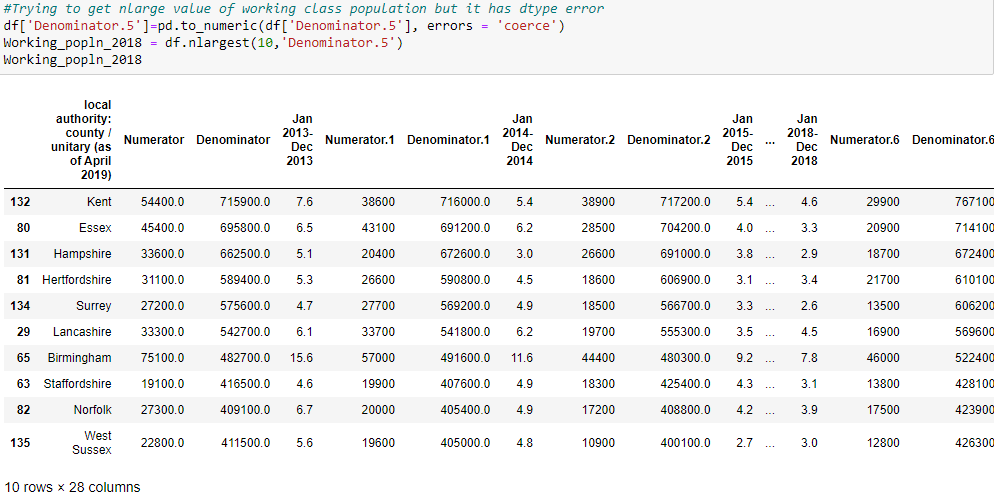
Description automatically generated









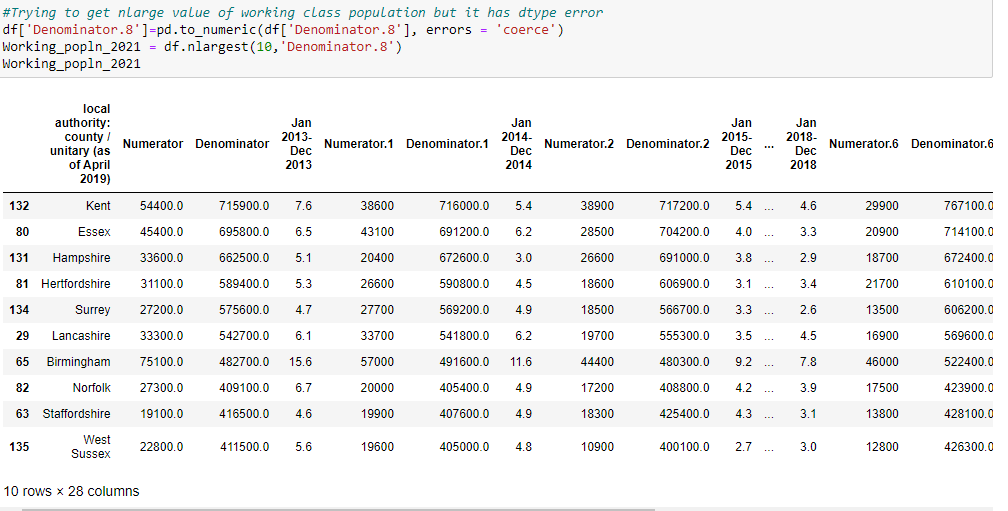


A picture containing table

Description automatically generated

Graphical user interface

Description automatically generated



**Ten Local Counties with Highest Unemployed Population**

The idea is to gather some insights regarding the highest working population corresponds to highest unemployment data. It will help me to compare the ten local counties of highest claimants with highest unemployed population. I am also trying to find out if the local counties with highest unemployed population will show the highest unemployment rate. The numerator in the column shows unemployed population.

Graphical user interface

Description automatically generated

Table

Description automatically generated

Table

Description automatically generated with low confidence

Graphical user interface, table

Description automatically generated with medium confidence

Graphical user interface, application

Description automatically generated

Table

Description automatically generated

Table

Description automatically generated

Table

Description automatically generated

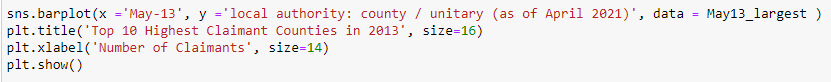
Table

Description automatically generated

**Exploratory Data Analysis**

* **Ten Local Authorities with Highest Claimants Count of Age above 16 years**

I used sns.barplot to visualise the highest claimants count of age 16+ from 2013-2022.



Chart, bar chart, funnel chart

Description automatically generatedChart, bar chart

Description automatically generated

Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generatedChart, bar chart

Description automatically generated

Chart, bar chart, funnel chart

Description automatically generated  
Chart, bar chart, funnel chart

Description automatically generatedChart, bar chart, funnel chart

Description automatically generated Chart, bar chart, funnel chart

Description automatically generated Chart, bar chart, funnel chart

Description automatically generated

Birmingham, Kent, Essex, Leeds, Glasgow City, Lancashire, and Manchester are seven local counties which appeared in top ten highest claimants count of above 16 years age throughout the last 10 years (2013-2022).

* **Ten Local Authorities with Highest Claimants as Proportion of Residents Aged 16-64 Years**

Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generated

Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generated

Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generated

Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generated

Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generated

* **Ten Local Counties with Highest Claimants of Age 25-49 Years**

Chart, bar chart

Description automatically generated Chart, bar chart

Description automatically generated

Chart, bar chart, funnel chart

Description automatically generatedChart, bar chart

Description automatically generated  
Chart, bar chart

Description automatically generated Chart, bar chart

Description automatically generated

Chart, bar chart, funnel chart

Description automatically generatedChart, bar chart, funnel chart

Description automatically generated

Chart, bar chart, funnel chart

Description automatically generated Chart, bar chart

Description automatically generated

* **Ten Local Authorities with Highest Claimants Aged 16-24 Years**

Chart, bar chart, funnel chart

Description automatically generated Chart, bar chart, funnel chart

Description automatically generated

Chart, bar chart

Description automatically generated Chart, bar chart, funnel chart

Description automatically generated

Chart, bar chart

Description automatically generatedChart, bar chart, funnel chart

Description automatically generated Chart, bar chart, funnel chart

Description automatically generated

Chart, bar chart

Description automatically generated

Chart, bar chart

Description automatically generated Chart, bar chart, funnel chart

Description automatically generated

* **Ten Local Authorities with Highest Claimants of Age Over 50 Years**

Chart, bar chart

Description automatically generated Chart, bar chart

Description automatically generated

Chart, bar chart

Description automatically generated Chart, bar chart

Description automatically generatedChart, bar chart, funnel chart

Description automatically generated Chart, bar chart, funnel chart

Description automatically generated

Chart, bar chart

Description automatically generatedChart, bar chart, funnel chart

Description automatically generated  
Chart, bar chart, funnel chart

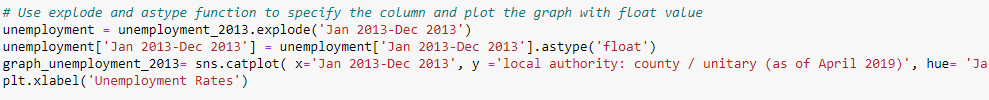
Description automatically generated Chart, bar chart

Description automatically generated

**Unemployment Data**

* **Ten Local Counties with Highest Unemployemnt rate**

I used astype function to plot float value through sns.catplot to visualise the unemplyement rate.



Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generatedApplication

Description automatically generated with low confidence

Graphical user interface, application

Description automatically generated with medium confidenceGraphical user interface

Description automatically generated with medium confidence

A picture containing graphical user interface

Description automatically generatedA picture containing text

Description automatically generated

Chart, scatter chart

Description automatically generated

* **Ten Local Counties with Largest Working Age Population and their Unemployment rate**

A picture containing graphical user interface

Description automatically generatedA picture containing chart

Description automatically generated

Chart

Description automatically generatedGraphical user interface

Description automatically generated with low confidence

A picture containing graphical user interface

Description automatically generatedChart

Description automatically generated

A picture containing graphical user interface

Description automatically generatedGraphical user interface

Description automatically generated with medium confidence

Graphical user interface

Description automatically generated with medium confidence

* **Ten Local Counties with Largest Unemployed Working Age Population**

Graphical user interface

Description automatically generated with medium confidenceA picture containing graphical user interface

Description automatically generated

A picture containing graphical user interface

Description automatically generatedGraphical user interface

Description automatically generated with low confidence

A picture containing graphical user interface

Description automatically generatedGraphical user interface

Description automatically generated with medium confidence

Graphical user interface

Description automatically generated Graphical user interface

Description automatically generated with medium confidence

Graphical user interface

Description automatically generated with medium confidence