

Lab Assignment – IA9 – Word Tree and Word Cloud

Note on Software used for following Visualizations: (Tableau)

Microsoft Excel is a spreadsheet developed by Microsoft for Windows, macOS, Android and iOS. It features calculation or computation capabilities, graphing tools, pivot tables, and a macro programming language called Visual Basic for Applications (VBA). Excel forms part of the Microsoft Office suite of software.

It includes:

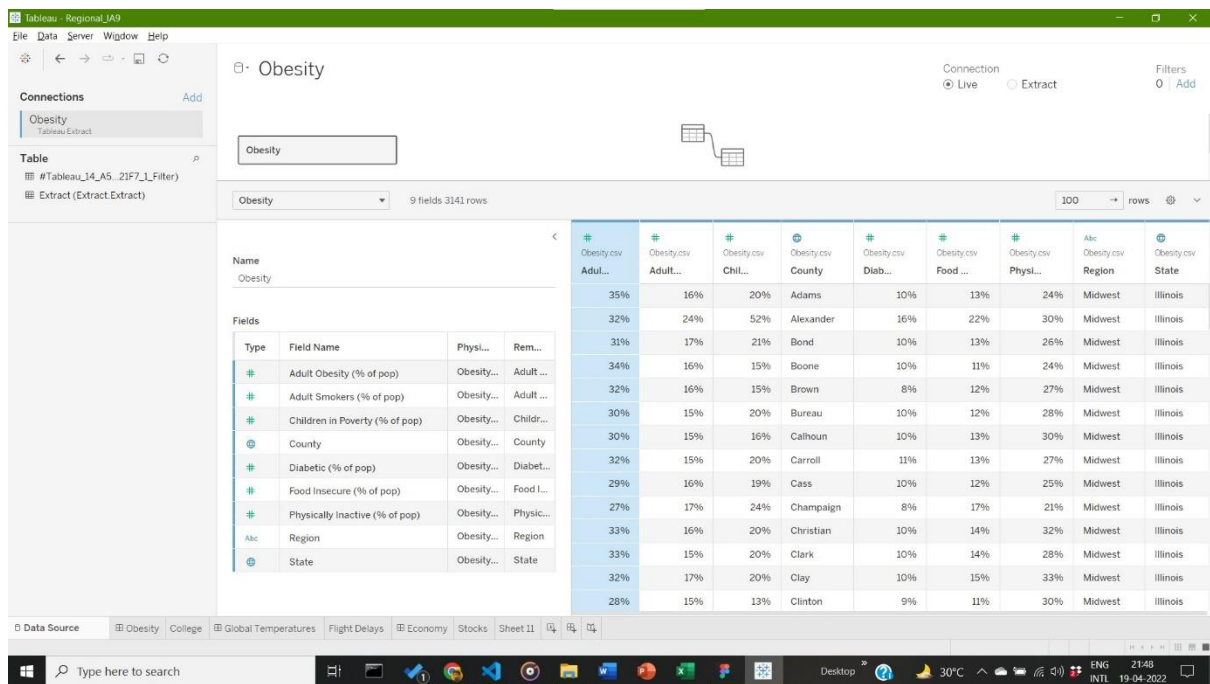
- Easy to access from different sources.
- No need for any technical or programming knowledge, and Quick response for making a dashboard.
- In terms of connecting and sharing, it has various inbuilt advanced features such as: Collaboration and distribution, highly securable, Multiple data sources connection, Easy importation and exportation of the massive size of data.
- For easy accessibility and analysis, the data file can be downloaded locally on mobile or desktop, multilingual representation of data, real-time exploration of any dataset, etc.

Q1) Using the following text visualization methods, create visualization of a text document of your interest: A: WordCloud

Answer: Some Key points to note before visualization process:

1. Dataset Used:

The dataset used is the inbuilt dataset **Regional (Obesity)** available in Tableau. The dataset depicts the obesity rates in all the states of the United States of America.

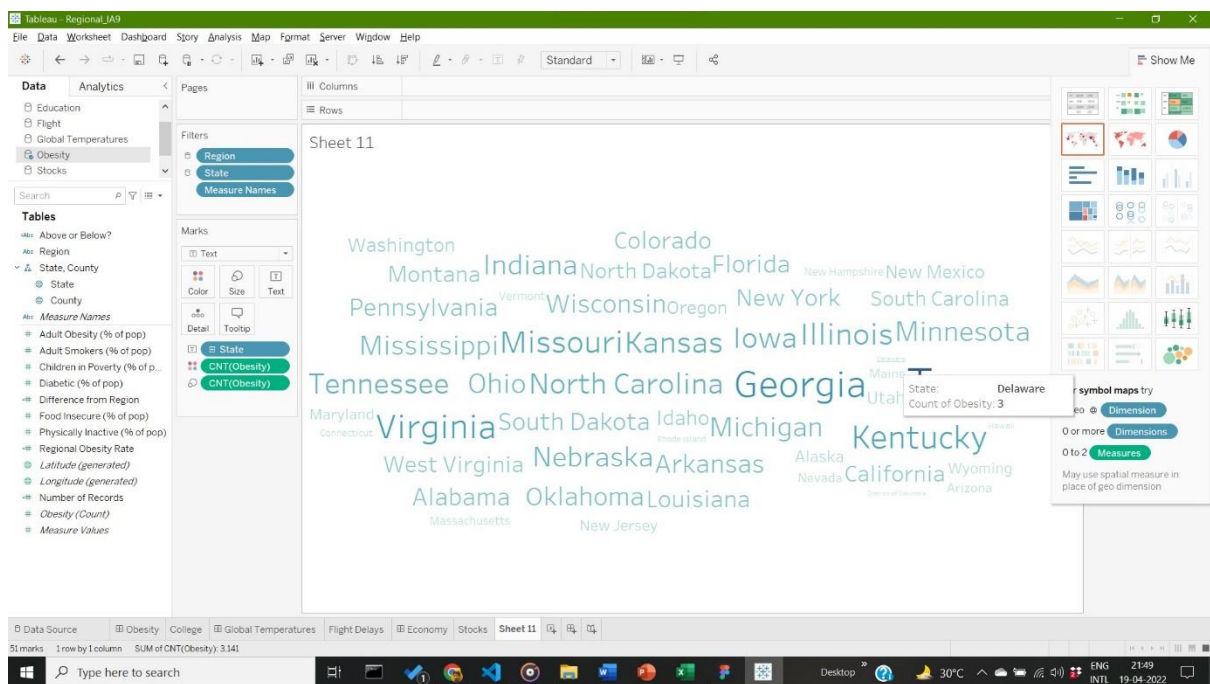
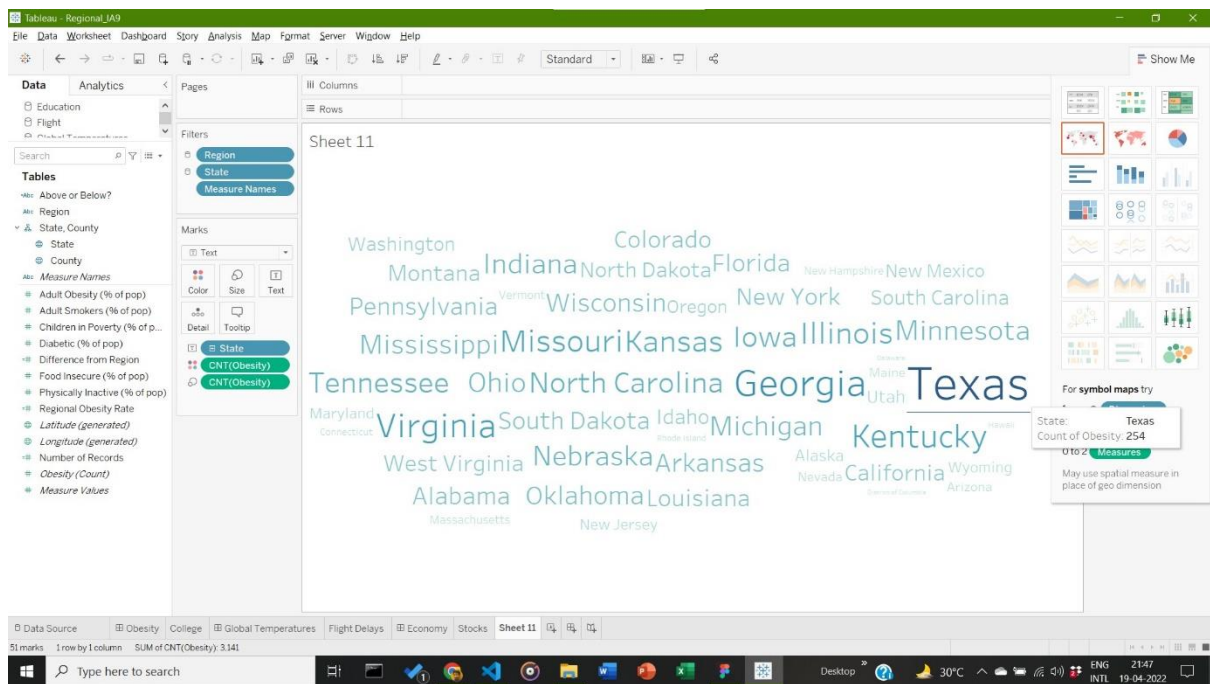


Fields

Type	Field Name	Physi...	Rem...
#	Adult Obesity (% of pop)	Obesity...	Adult ...
#	Adult Smokers (% of pop)	Obesity...	Adult ...
#	Children in Poverty (% of pop)	Obesity...	Childr...
🌐	County	Obesity...	County
#	Diabetic (% of pop)	Obesity...	Diabet...
#	Food Insecure (% of pop)	Obesity...	Food I...
#	Physically Inactive (% of pop)	Obesity...	Physic...
Abc	Region	Obesity...	Region
🌐	State	Obesity...	State

2. Visualization:

Visualization based on States available (denoted in terms of text), Count of Obesity Cases (Denoted in two ways: 1st depicts count via colour and 2nd depicts count via size). With overall marks of type Text. Final Results are as shown below:



Inference:

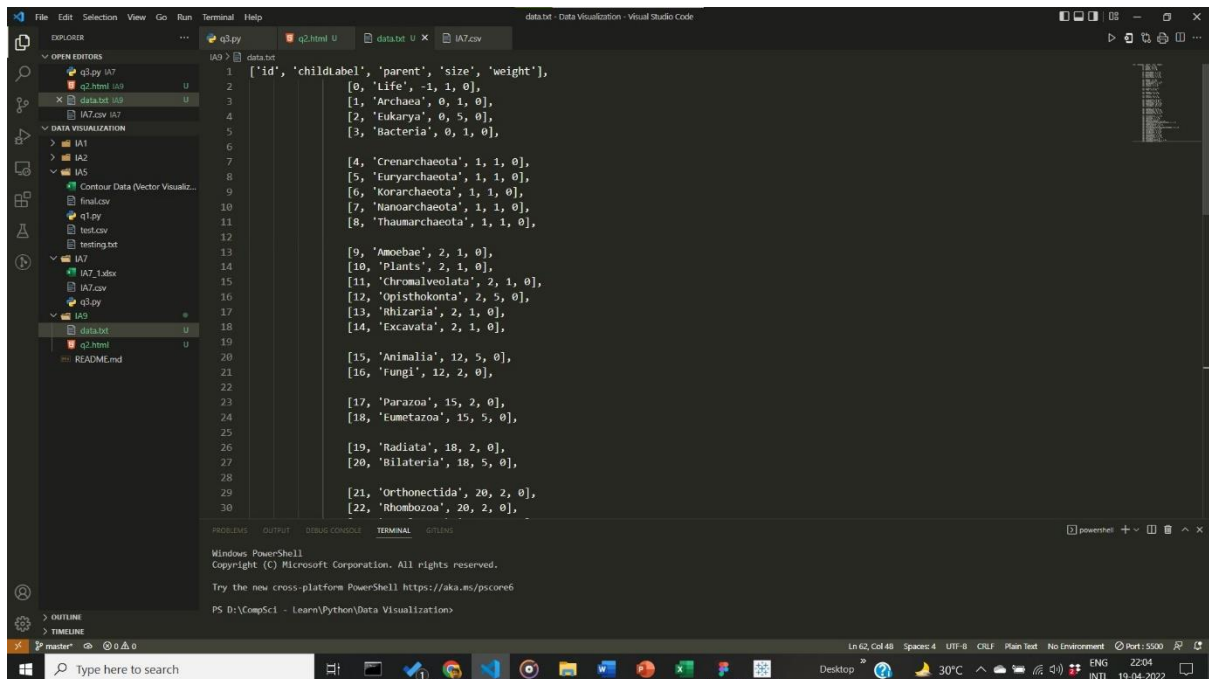
The State with highest Obesity count is Texas with a total count of **254**, closely followed by Georgia; whereas the state with the lowest count is Delaware with a total count of **3**.

Q2) Using the following text visualization methods, create visualization of a text document of your interest: A: WordTree

Answer: Some Key points to note before visualization process:

1. Dataset Used:

The dataset used is the classification of Micro organisms' dataset, currently available in the given link: <https://developers.google.com/chart/interactive/docs/gallery/wordtree>. It depicts the hierarchy found in the classification of micro-organism life. (shown here in data.txt)



The screenshot shows the Visual Studio Code interface with a file named 'data.txt' open. The file contains a JSON array of objects, each representing a node in a hierarchical tree structure. The objects are formatted as follows:

```
[{"id": "Life", "parent": null, "size": 1, "weight": 1}, {"id": "Archaea", "parent": "Life", "size": 1, "weight": 1}, {"id": "Bacteria", "parent": "Life", "size": 1, "weight": 1}, {"id": "Eukarya", "parent": "Life", "size": 1, "weight": 1}, {"id": "Crenarchaeota", "parent": "Archaea", "size": 1, "weight": 1}, {"id": "Euryarchaeota", "parent": "Archaea", "size": 1, "weight": 1}, {"id": "Korarchaeota", "parent": "Archaea", "size": 1, "weight": 1}, {"id": "Nanoarchaeota", "parent": "Archaea", "size": 1, "weight": 1}, {"id": "Thaumarchaeota", "parent": "Archaea", "size": 1, "weight": 1}, {"id": "Amoebozoa", "parent": "Eukarya", "size": 2, "weight": 1}, {"id": "Plants", "parent": "Eukarya", "size": 2, "weight": 1}, {"id": "Chromalveolata", "parent": "Eukarya", "size": 2, "weight": 1}, {"id": "Opisthokonta", "parent": "Eukarya", "size": 2, "weight": 1}, {"id": "Rhizaria", "parent": "Opisthokonta", "size": 2, "weight": 1}, {"id": "Excavata", "parent": "Opisthokonta", "size": 2, "weight": 1}, {"id": "Animalia", "parent": "Opisthokonta", "size": 12, "weight": 5}, {"id": "Fungi", "parent": "Opisthokonta", "size": 12, "weight": 2}, {"id": "Parazoa", "parent": "Opisthokonta", "size": 15, "weight": 2}, {"id": "Eumetazoa", "parent": "Opisthokonta", "size": 15, "weight": 5}, {"id": "Radiata", "parent": "Eumetazoa", "size": 18, "weight": 2}, {"id": "Bilateria", "parent": "Eumetazoa", "size": 18, "weight": 5}, {"id": "Orthonectida", "parent": "Bilateria", "size": 20, "weight": 2}, {"id": "Rhombozoa", "parent": "Bilateria", "size": 20, "weight": 2}],
```

2. Visualization:

Visualization based on google.charts.load[wordtree] package, made available by the Google dev Society. It is formatted to display data explicitly with its suffix brought into prominence. Word Tree Constructed via HTML, CSS and Javascript as given library is a JS library package.

```

1  <html>
2
3  <head>
4
5    <script type="text/javascript" src="https://www.gstatic.com/charts/loader.js"></script>
6    <script type="text/javascript">
7      google.charts.load('current', { packages: ['wordtree'] });
8      google.charts.setOnLoadCallback(drawSimpleNodeChart);
9      function drawSimpleNodeChart() {
10        var nodeListData = new google.visualization.arrayToDataTable([
11          ['id', 'childLabel', 'parent', 'size', 'weight'],
12          [0, 'Life', -1, 1, 0],
13          [1, 'Archaea', 0, 1, 0],
14          [2, 'Eukarya', 0, 5, 0],
15          [3, 'Bacteria', 0, 1, 0],
16
17          [4, 'Crenarchaeota', 1, 1, 0],
18          [5, 'Euryarchaeota', 1, 1, 0],
19          [6, 'Korarchaeota', 1, 1, 0],
20          [7, 'Nanoarchaeota', 1, 1, 0],
21          [8, 'Thaumarchaeota', 1, 1, 0],
22
23          [9, 'Amoeboae', 2, 1, 0],
24          [10, 'Plants', 2, 1, 0],
25          [11, 'Chromalveolata', 2, 1, 0],
26          [12, 'Opisthokonta', 2, 5, 0],
27          [13, 'Rhizaria', 2, 1, 0],
28          [14, 'Excavata', 2, 1, 0],
29
30          [15, 'Animalia', 12, 5, 0],
31          [16, 'Fungi', 12, 2, 0],

```

