**Budget Text Processing**

1. **Introduction**

In this project we will be analyzing Adopted budget text from the different counties of the State of North Carolina. The ‘Adopted Budget Plan’ is the **annual budget approved by the Board of Supervisors for the fiscal year which runs from July 1 through June 30**. The scope of this project is Adopted Budgets for FY 2019-2020 each.

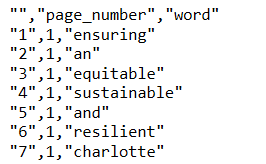
1. **Data:**

The budget texts will be fetched from the following counties or cities as mentioned below:

* City of Charlotte
* Mecklenburg County
* Wake County
* City of Raleigh
* Guilford County
* City of Durham
* Durham County

The Budget documents (.pdf) obtained from respective counties/cities is converted to two types of csv files as shown below for City of Charlotte:

1. Simple tokenization



1. Emotion categorization

