Daily Log

Monday November 18

Began working with the Text Summarization API TextTeaser. Read through README.md File on Github (Realized the Importance of One) and Began Downloading Appropriate Packages.

Tuesday November 19

Continued Working on Fixing Package Errors with TextTeaser. However, when Errors Stopped Occurring, Blank Texts Returned as the Summary. Looked into Potential Reasons why and Some Resources said Periods/Punctuation were Required for Accurate Text Summarization. Decided to Work with Self Created Text, but TextTeaser still Did not Work.

Thursday November 21

Looked back at https://towardsdatascience.com/text-summarization-in-python-76c0a41f0dc4 to Look Into Other Potential APIs. Gensim's Text Summarization Seemed to have Potential to Yield Good Results, so Began Downloading Appropriate Packages. Started to Implement Code using Self Created Text. Noticed that Commas would Cause Inaccuracies in Summarization.

Timeline

Date	Goal	Met
Nov 8	Research Possible Text Summarization APIs and Methods	Yes, Read Different Articles about Ways to Summarize Text
Nov 15	Continue to Research Possible Text Summarization APIs and Methods	Yes, Decided to Move Forward with TextTeaser
Nov 22	Begin Implementing a Text Summarization Algorithm	Yes, Started Working with APIs (Text Teaser and Gensim)
Nov 29	Continue Implementing a Text Summarization Algorithm	
Dec 6	Finish Implementing a Text Summarization Algorithm for Piece of Text	
Winter Goal	Summarize a TED Talk by Implementing a Text-Summarization API	

Reflection

It was great to get back to coding and begin implementing APIs. However, it was frustrating to spend a majority of my time on downloading packages and troubleshooting import errors. I was pleased to finally find results with Gensim, even though they were not optimal. It seemed that periods in text were better for text summarization, which is an issue as the IBM Watson Speech to Text translations did not add punctuation. However, when using the language English as a parameter the translations added "HESITATION" to the transcript. If punctuation is truly needed I plan to go back and replace "HESITATION" with period, as it is a likely sign that is the end of a sentence as most speakers usually pause. Next week I plan to continue working with Gensim and hopefully yield better results.