Project Question/Idea:

Pancreatic cancer is one of the deadliest cancers in terms of long term survival as compared to the number who are diagnosed. Around 91% of patients diagnosed, will not leave 5 years past diagnosis. This is due to a lack of treatments as well as a lack of ability to diagnose the disease adequately. The goal of this project is to apply a temporal evolutionary approach to predict the sequence of mutations in Pancreatic Adenocarcinoma. This could be used to determine future therapeutic targets as well as intervention points and diagnostic markers for tumorigenesis. In addition, the second goal is to subtype these mutations as related to colorectal cancers to determine possible therapy directions as these two cancers often share similar gene mutations.

Methodology:

The data used in this project will be from the NIH GDC Data Portal. The method used will be Retracing the Evolutionary Steps in Cancer (RESIC). The first step is using GISTIC to identify genes targeted by somatic copy-number alterations (copy number variant) that drive cancer growth. These are then ranked by their pairwise association using Fischer’s Exact Test. Then, the most likely sequence event is identified by RESIC. The results from this RESIC are used to reconstruct the order in which these alterations arise.

To complete the second goal, the algorithm of Maximum Likelihood will be used as this best fits the copy number variant (CNV) data that will be produced from the first part of this project.

Git Repository:

https://github.com/MonicaBrown28/finalproject