

List of Team Members: Noah Stapp, Richa Gadgil

Initial Decisions:

We decided to go with Python, using Pandas. Python is used extensively for data analysis and its simple IO and language structure made it an optimal choice. Pandas is a very powerful library that makes parsing and searching data extremely easy.

Notes on selected internal architecture:

We used a Pandas dataframe to store all of the contents of the students.txt file. We were able to easily manipulate this data frame in order to get the statistics we needed. A dataframe represents a relational model. Each command simply queried the dataframe and filtered the results to retrieve the desired data.

Task log:

Project ideas and brainstorm: 10:10-10:20 [Richa]

Parsing txt file and creating a Pandas dataframe: 10:20-10:30 [Noah]

Created first few query functions: 10:30-11:00 [Richa]

Implemented query functions according to spec: 10:10-11:00 [Richa and Noah]

Testing: 11:30-12:30 [Richa and Noah]

Writeup: Richa & Noah

Notes on testing. How many bugs found, how long it took to find them:

Several bugs were found while testing, usually as a result of not verifying valid or correct input. One common bug we found was that we forgot to verify the existence of keys before searching the data frame, this led to a lot of errors we were able to fix. We also fixed errors in prompt control flow to make them conform to the spec. Finding and fixing all bugs did not take very long, as most were easily fixed by checking input before calling functions.