Project 2 Summary

Alireza Ahmadinia

During the second phase of the software project, I chose to focus on the Pandas library due to its extensive functionality, Python foundation, and a wide range of open issues. After exploring multiple issues, I selected issue #60567, which involved challenges with the arrowDtype and the groupby function. The complexity of this issue required an in-depth analysis of various Pandas modules, including dtypes, dataframe_protocol, and groupby files . Throughout this process, I encountered and overcame significant challenges, such as navigating a large-scale codebase, understanding others' coding styles, and dealing with a lack of structural diagrams for the code. To address this, I created detailed diagrams illustrating the relationships between classes and modules.

- Issue

- Pull Request

- Groupby Diagram

- Dtypes Diagram