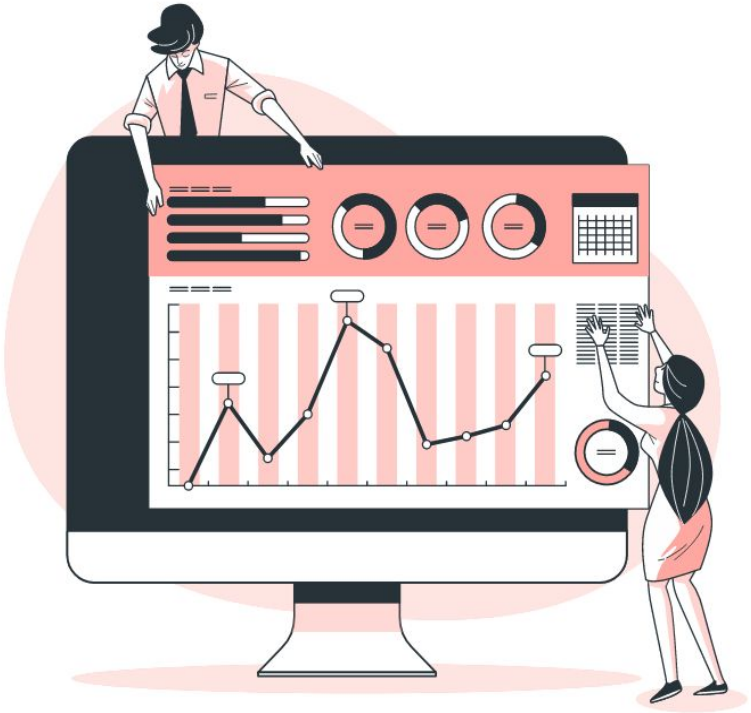
MASSACHUSETTS ANALYTICS

x

Capstone Team A7



Our team

Qianru (Alice) Ai

Keshuo Liu

Ana Maharjan

Mekhal Raj

Zheming Xu

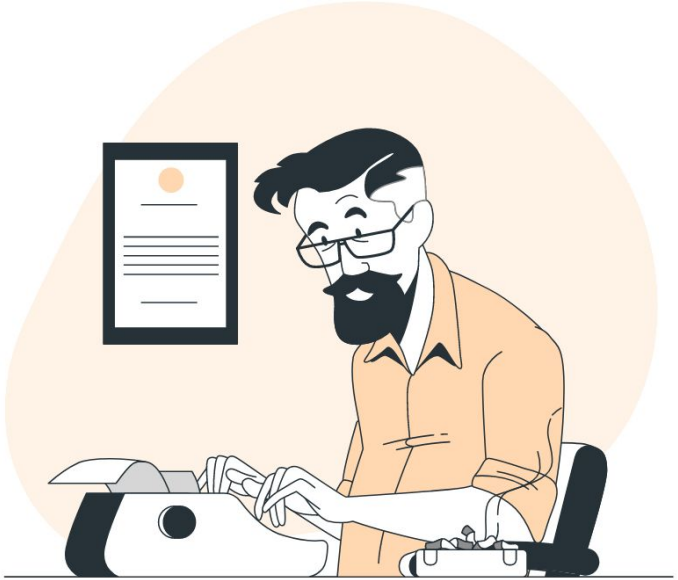
Table of contents

01 Project Overview

02 App Design & Project Roadmap

03 Data Analysis

04 Conclusion



01
Project Overview

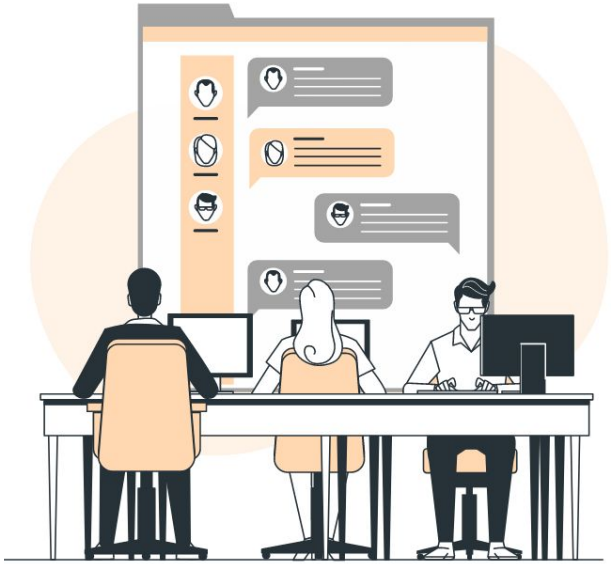
Business Context: Key Terms

What is SEC?

U.S. government oversight agency responsible for regulating the securities markets and protecting investors.

MMF Data:

Money market funds are mutual funds that invests in highly liquid, short-term assets (i.e. cash, cash equivalent securities). These funds are registered and regulated under the SEC.



Project Objective

Develop an application to scrape, process & clean MMF data from SEC.gov while providing valuable insights on historial holding patterns.

Project Benefits

Leverage Fast-Growing Financial Data

Simple and Convenient Method to Access, Download & Process Data

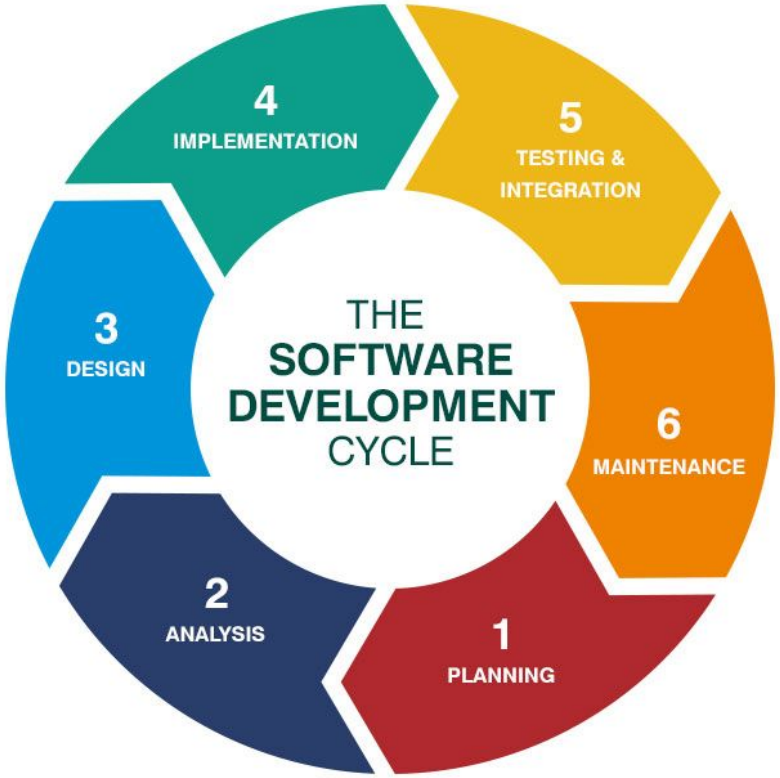
Analyze Past Investor Holding Patterns

Provide Client with Tools to Benefit their Business Model & Goals

02

App Design

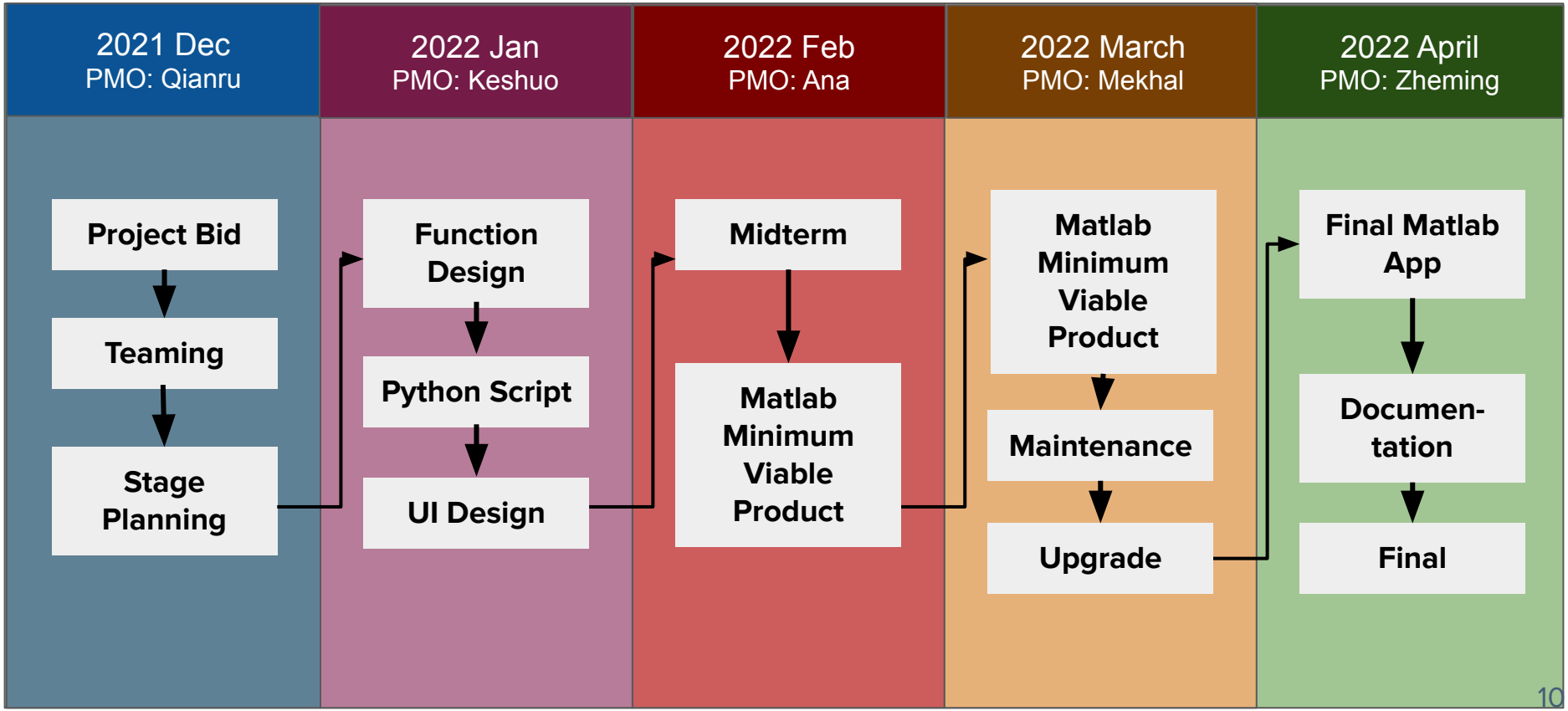
Software Development Life Cycle



1. Planning (workflow)
2. Analysis (main requirement)
3. Design (key functions)
4. Implementation (Python & Matlab Programming)
5. Testing & Integration

   (Minimum Viable Product)
6. Maintenance (Debug)

Planning & Workflow



Requirement Analysis

Users

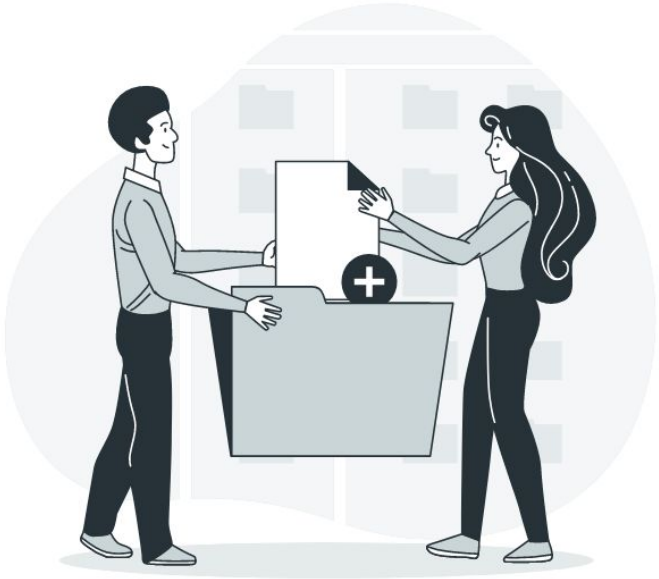
Massachusetts Analytics' employees

Preferred Platform

Matlab

Main Webpage

Company search page of SEC.gov EDGAR database



Key Functions

- Web scraping
- Processing XML files
- Filters
- Output CSV files

Custom Features

- User-friendly
- Can choose file saving directory
- By individual or by group
- Charts with customized colormap

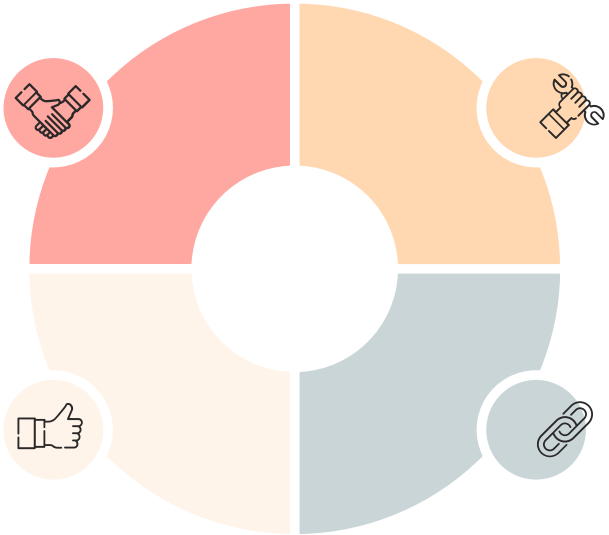
Design & Key Functions

Web Scraping

https://www.sec.gov/cgi-bin/browse-edgar?action=getcompany&CIK=**0001217286**&type=N-MFP2&dateb=&owner=include&start=0&count=100

Filters

Date, Fund, Series, Classes, Security Holdings



Processing XML Files

```
<formData>
    <generalInfo>...</generalInfo>
    <scheduleOfPortfolioSecuritiesInfo>
        <repurchaseAgreement>...</repurchaseAgreement>
    </scheduleOfPortfolioSecuritiesInfo>
</form Data>
```

Output CSV Files

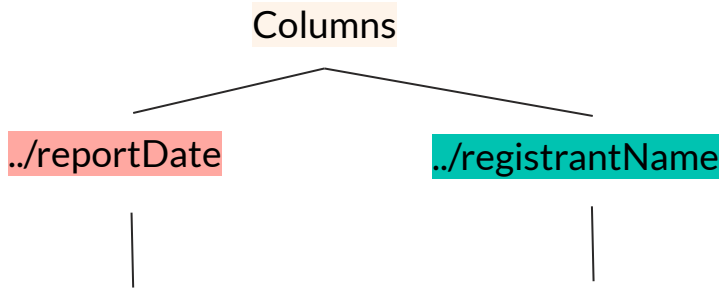
Organize multi-level data to 1- or 2-dimensional table

Implementation: Python VS. Matlab

	Web Scraping	Read XML Files	Batch Download	Output CSV File	UI
Python	▪ Urllib.request ▪ BeautifulSoup (html.parser)	▪ BeautifulSoup (lxml) ▪ lxml.etree ▪ xml.etree.ElementTree	▪ Multiprocessing	▪ Pandas	▪ Figma ▪ Google Slides
Matlab	▪ websave	▪ readtable (detectImportOptions)	-	▪ writetable	▪ Matlab App ▪ Designer

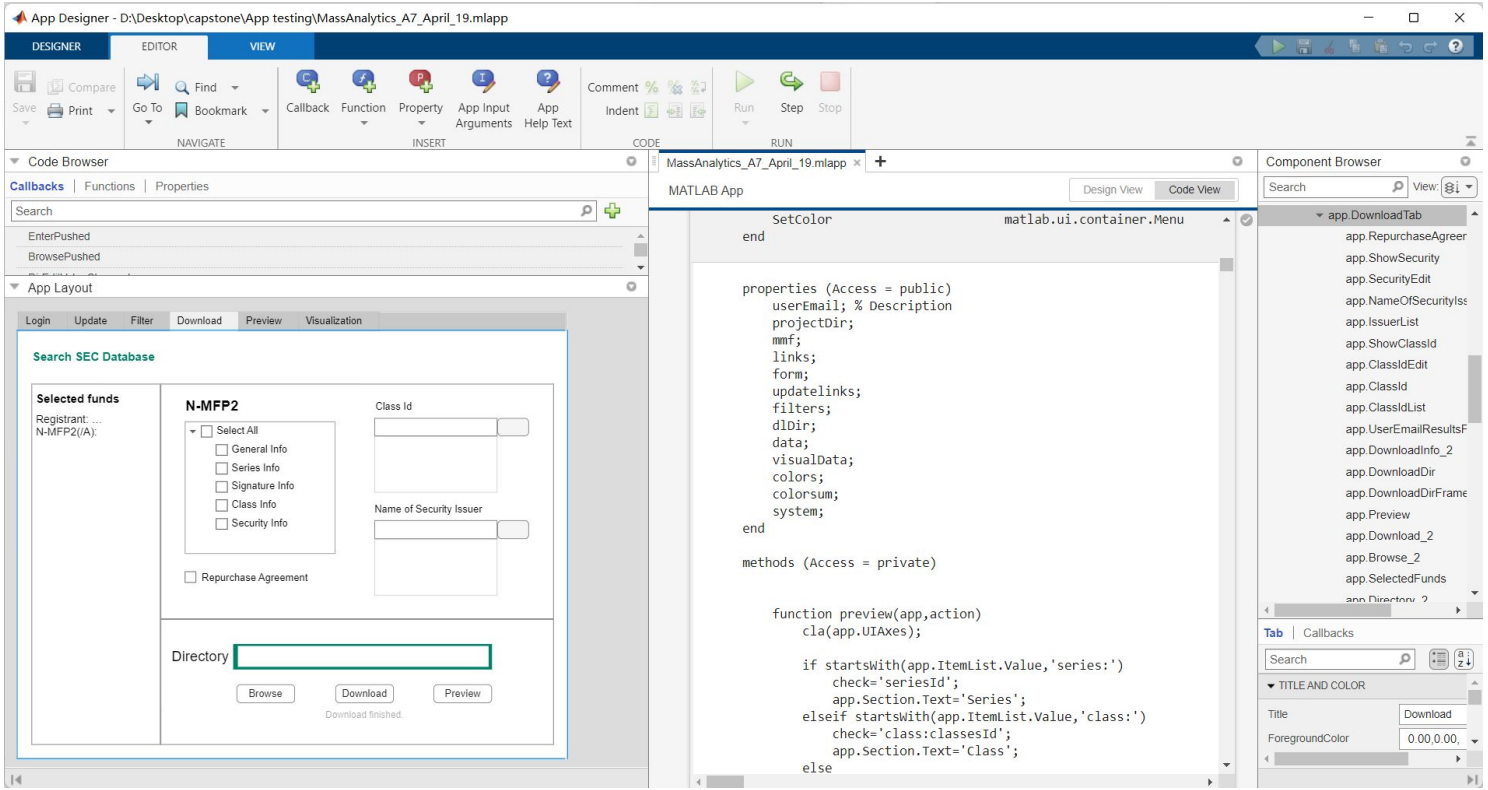
Python VS. Matlab

```
<formData>
    <generalInfo>
        <reportDate>
            2021-01-31
        </reportDate>
        <registrantName>
            JPMorgan Trust I
        </registrantName>
    </generalInfo>
</formData>
```

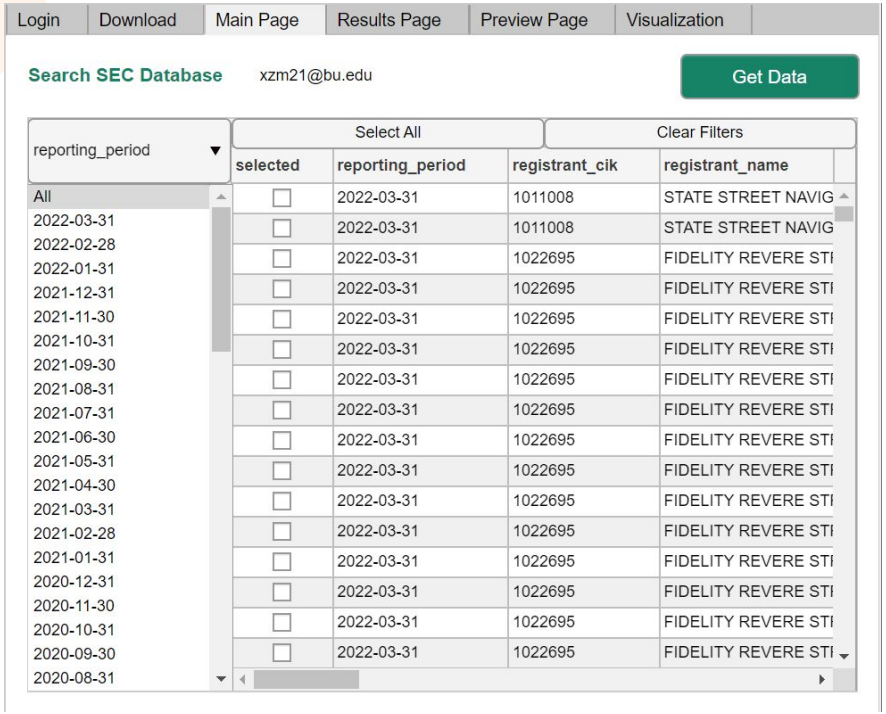


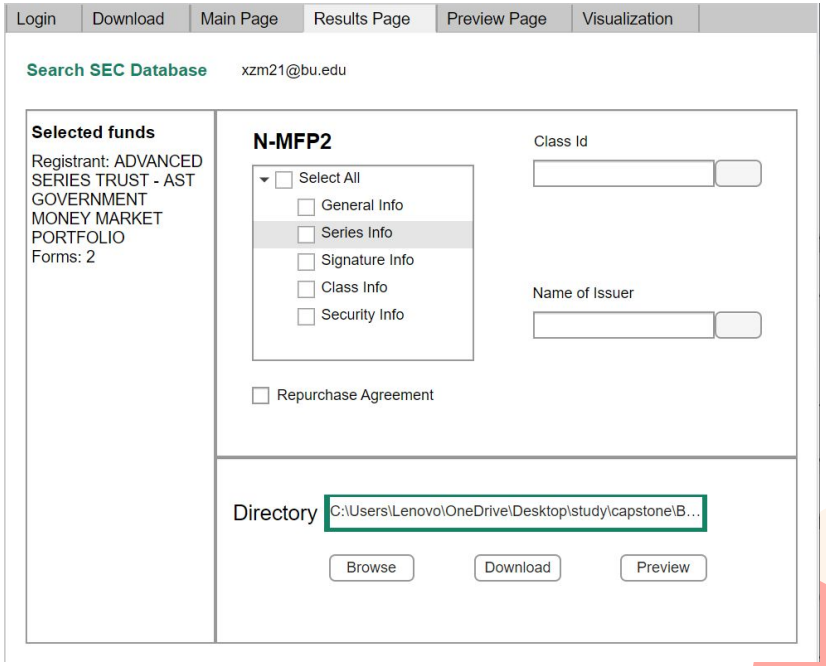
2021-01-31	JPMorgan Trust I
XXXX-XX-XX	XXX

Testing & Integration

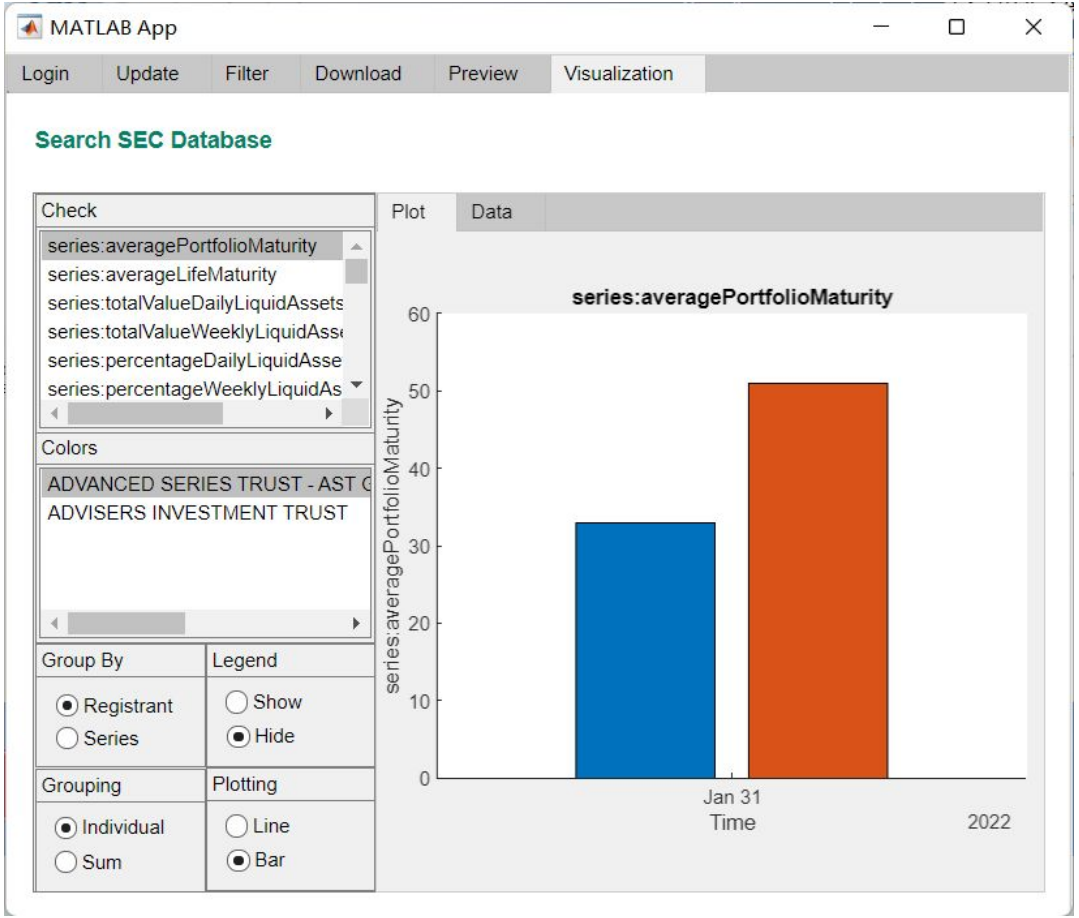


Testing & Integration





Testing & Integration



Maintenance & Debug

1

Connection

Website server download speed limit

2

Filters

Get holdings of funds for certain securities

3

Database

Mega size, and time-consuming for initial download

Maintenance (Solutions)

01

Connection

Set up a threshold and a batch size=100 for progress saving.

Users may restart the app to continue.

02

Filters

Choose forms.

Filter by class id or name of security issers.

03

Database

Loop over all forms to create an index including class id, series id, and name of security issuers.

03

Data Analysis

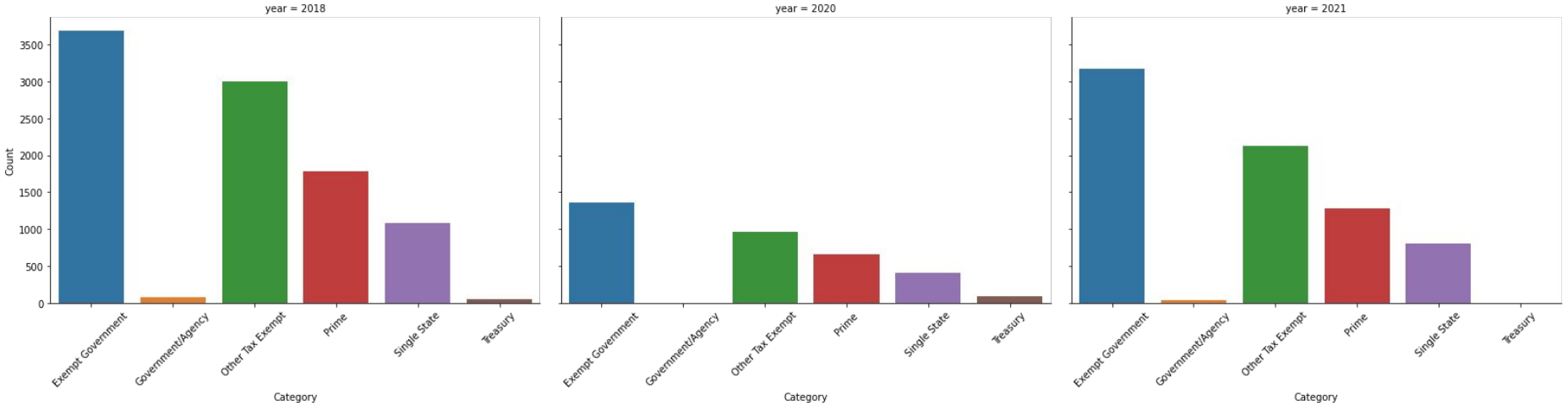
Objective

How does short term crisis impact the market?

- Change in series
  - Kind & quantity, yield, maturity
- Change in market
  - Yield, maturity
  - associate with interest rates

Series - Kind & quantity

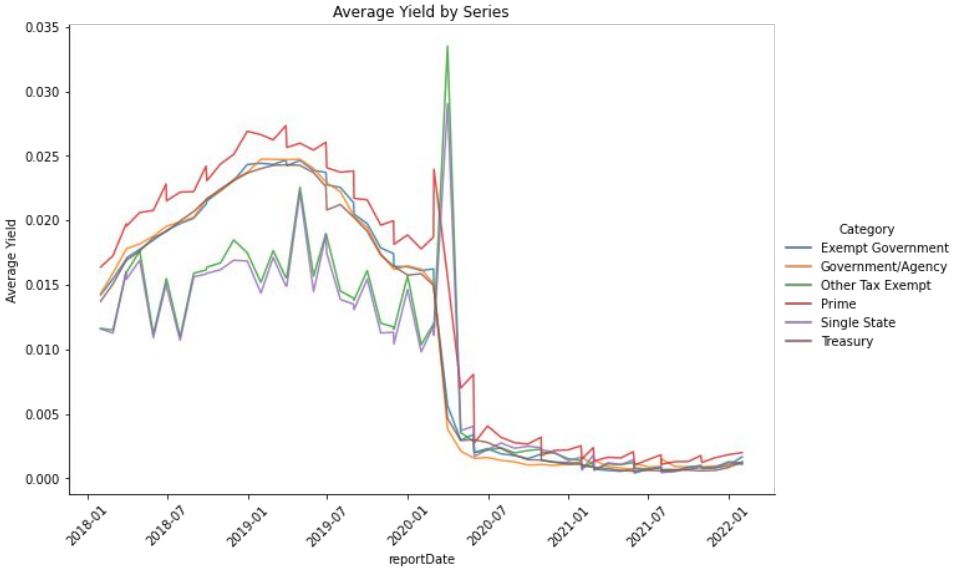
2018-01-31 (Pre pandemic), 2020-02-28 (Post pandemic) and 2021-02-28



Series - Yield

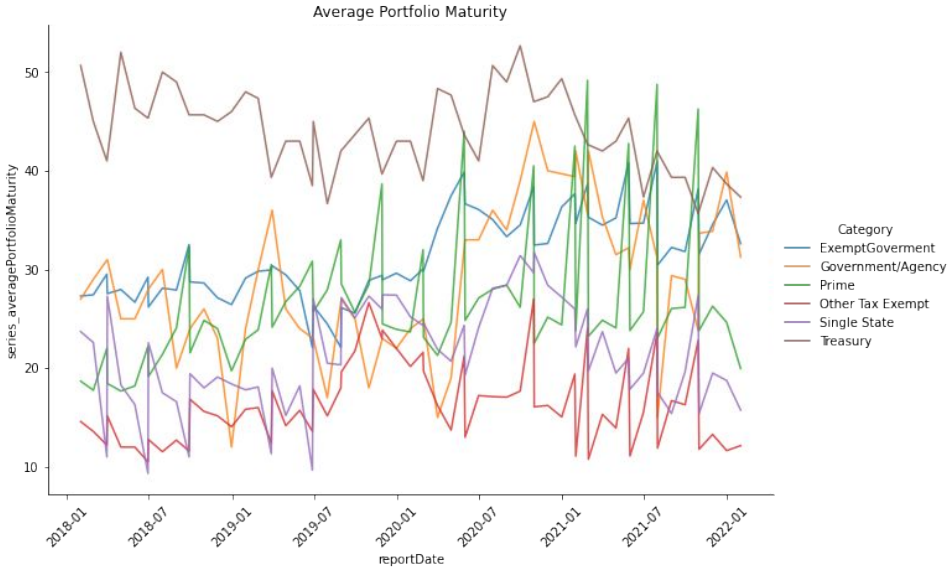
- Average return ranking: Prime, > Government/Agency ≈ Exempt government ≈ Treasury > Other tax exempt > single state

- Overall downward trend in yields.
- After the sharp decline following the peak in April 2020, the gap in yields was narrowed significantly.



Series - Portfolio Maturity

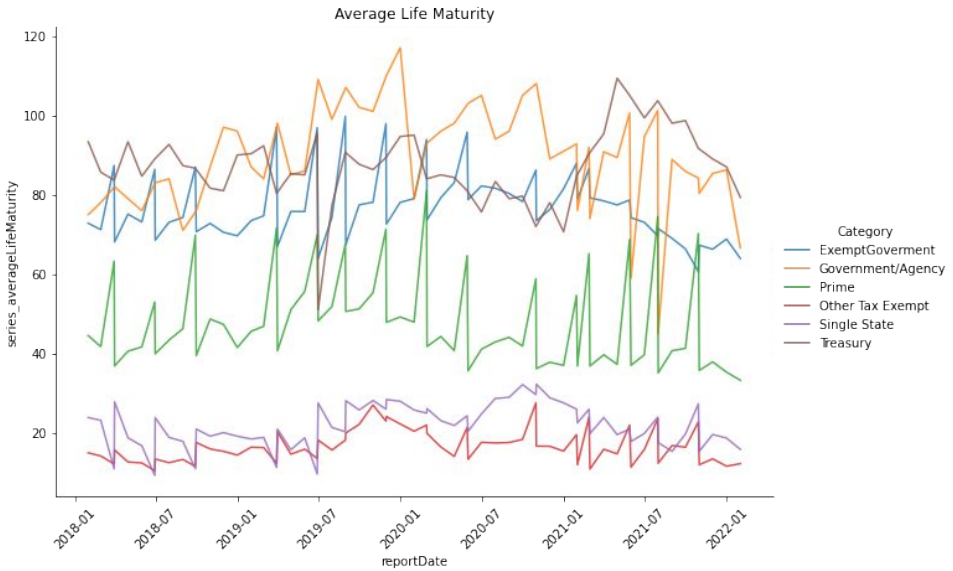
- October 2020:
  Coronavirus + Pre-election



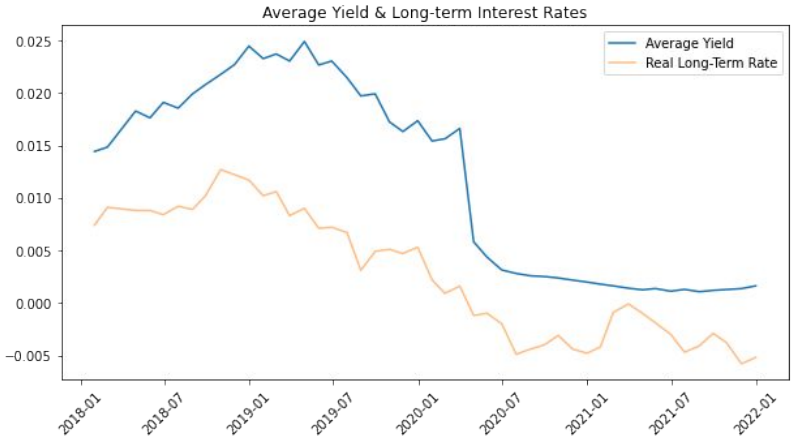
Series - Life Maturity

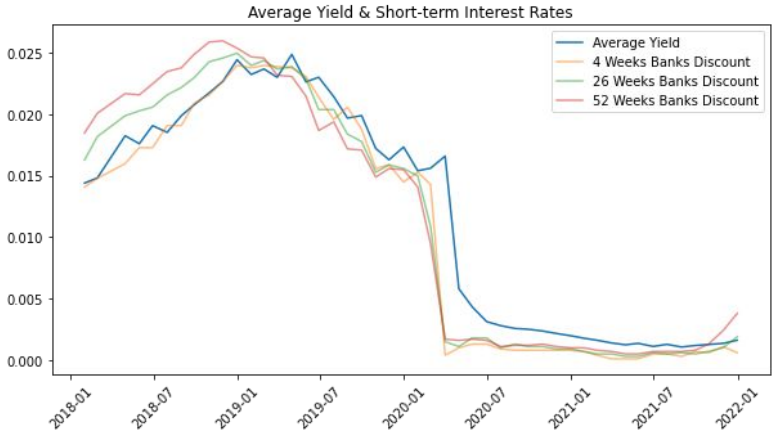
- Generally, life maturity of each category are relatively stable.

- Prime decreased.
- Government/agency & Exempt Government started decreasing after around Oct. 2020.
- Treasury gradually decreased during early pandemic, but increased sharply after around Jan. 2021.



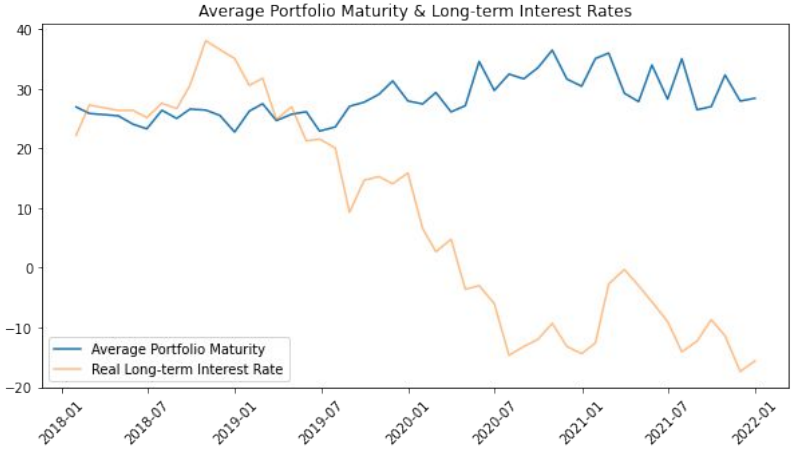
Market - Yield

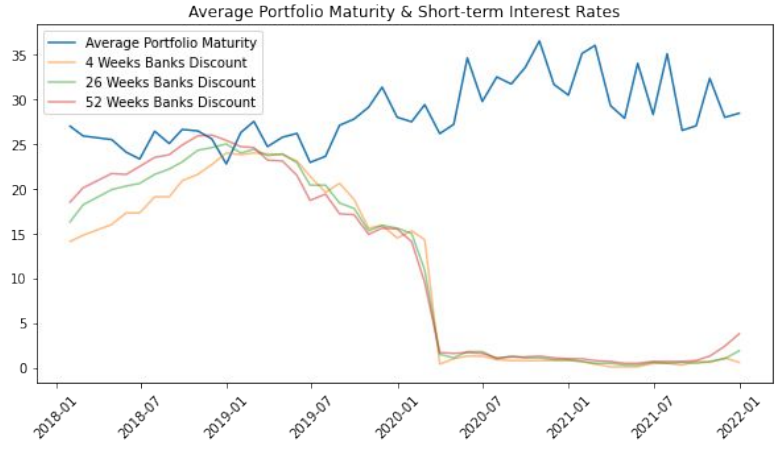




- The market yield follows the trend of long & short term interest rates.

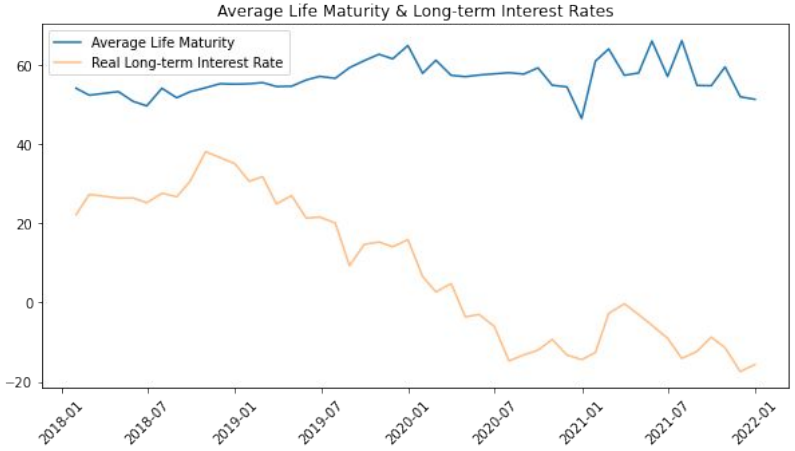
Market - Portfolio Maturity

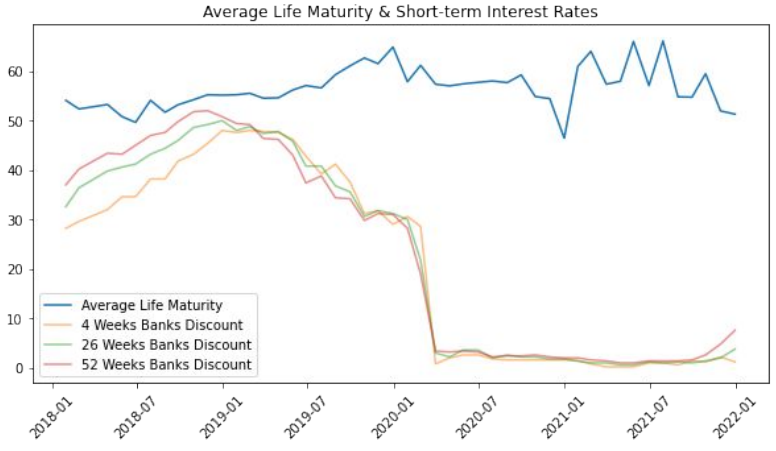




- The market average portfolio life is slightly negative to the interest rates' trends, but not so sensitive.

Market - Life Maturity

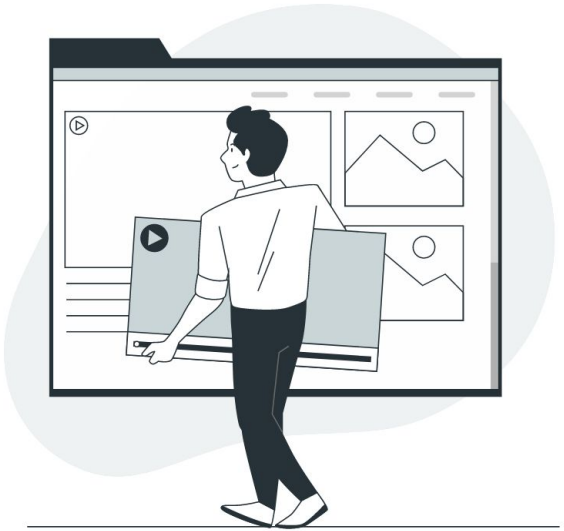




- The market average life maturity seems to be independent from long-term and short-term interest rates.

Conclusion & Insights

04



Important Analysis Findings

- When a market crisis arises, the U.S. government issues more Treasury debt to obtain sufficient funds to manage the crisis and stimulate market investment.
- Market turmoil keeps yields lower for all types of securities and can narrow the gap between yields for them.
- During crisis, investors prefer long-term investments due to their decline in confidence in the market.
- The market average yield follows the interest rates' trends, but the market average portfolio maturity and average life maturity are more independent from the interest rates' trends.

Project Highlights

- Created the web scraping application on Matlab.
- Used UI design techniques to improve the aesthetics and ease of use of the application.
- Exported MMF data since October 2016 with their monthly class and series information using the app.
- Generated analysis graphs for individual funds or grouped data.
- Troubleshooted and resolved post-production issues.

05

Challenges

Challenges

- Learning Matlab for creating the application.
- Accommodate differences between the Python language & packages and the Matlab language & variable types.
- Troubleshooting and fixing application errors - error due to no response from the sec.gov server.
- Improving efficiency of scraping and restructuring of large amounts of xml-format data.

Next Steps

06

Next Steps

1. Expanding the analysis reports by adding more predictive models to see if we can:

- Estimate the future asset class mix of investor holdings.
- Determine an investor's market sentiment from past investment patterns.

2. Improve App UI and incorporate more real-time visualization models.

Thanks!
Questions?